Dropping the rating curve: calibrating a rainfall-runoff model on stage to reduce discharge uncertainty

M. M. Piet

Delft University of Technology

TU Delft Delft University of Technology

Dropping the rating curve: calibrating a rainfall-runoff model on stage to reduce discharge uncertainty

By

M. M. Piet

For the degree of Master of Science in Civil Engineering
at Delft University of Technology,
to be defended publicly on Friday March 7, 2014 at 1:00 PM.

Committee

Prof. dr. ir. Hubert H.G. Savenije	Delft University of Technology
Prof. dr. ir. Matthijs Kok	HKV/Delft University of Technology
Dr. ir. Markus Hrachowitz	Delft University of Technology
Ir. Tanja Euser	Delft University of Technology
Hongkai Gao, MSc	Delft University of Technology

Summary

Direct measurement of river discharge is time consuming and financially demanding. Continuous river discharge measurements are therefore generally derived from continuous stage measurements, through a stage-discharge relation, also called a rating curve. Rating curves are determined by fitting a curve to a limited number of points (h_i, Q_i), whereby h_i and Q_i represent stage and discharge measured in a certain cross-section of a river at a fixed geographical location. Commonly these points (h_i, Q_i) originate from measurements done under regular flow conditions, due to which a considerable part of the curve is based on interpolation and extrapolation. Therefore the uncertainty in the rating curve particularly during floods can be considerable, which directly translates into uncertainty in the discharge data. Calibrating a model on uncertain discharge observations leads to biased model parameter estimates, which directly lead to biased model predictions.

This research shows an approach whereby a conceptual rainfall-runoff model is calibrated on the basis of stage data only. In addition to the existing conceptual model parameters, extra parameters have been added that define the rating curve. A stepwise calibration method has been applied whereby first the rating curve parameters were determined and subsequently the remaining model parameters. Once the rating curve parameters had been fixed, the reanalysed hydrograph was established using the observed stage readings. Subsequently, the model hypotheses on catchment behaviour were tested by conventional methods.

In this research these methods have been applied to the scarcely gauged Endau River catchment, located in the South-East of peninsula Malaysia. The initial results are promising. It was found that the rating curve parameters are well defined and optimise to values that correspond to the physical property they represent. When comparing the reanalysed rating curve with the original rating curve it can be seen that the initial part of the rating curve overlaps, corresponding with the most reliable part of the original rating curve. When comparing the reanalysed rating curves with discharge measurements, a high correspondence is observed, while the similarity between the measurements and the original rating curve is very low. Finally, the modelled hydrographs appear to be relatively well able to mimic the reanalysed hydrograph, even though it was impossible to find a proper model for the original hydrograph.

To test the sensitivity of the calibrated rating curve to model structure, both a lumped and a topography driven model structure have been tested. Additionally, these models were exposed to two rating curve definitions and a variable model forcing. The results of this sensitivity analysis show that the calibrated rating curve is relatively insensitive to model structure and relatively sensitive to the number of parameters used for the rating curve definition. Concerning the forcing, the rating curve is highly sensitive to precipitation and less sensitive to potential evaporation. The reason for this sensitivity is that the established rating curve is strongly determined by the water balance, which is dominated by the primary driver P.

The calibrated models have been validated on independent data by split-sample validation and by transfer to a different catchment. The topography driven models appear to perform best during both forms of validation. An additional way of validating was carried out by coupling the best performing rainfall-runoff model to a steady state salt intrusion model: a novel approach which offers mutual model validation. The results show that the predicted discharge of the rainfall-runoff model corresponds rather well with the discharge determined by the salt model.

So, the most important conclusions from this research are: calibration of a rainfall-runoff model on stage reduces discharge uncertainty in scarcely gauged basins; topography driven models are better transferable than lumped models; and linking a salt intrusion model to a rainfall-runoff model offers mutual model improvement.

Acknowledgements

This report is the final product of my graduation project, which I could never have written without the support of many people.

Foremost, I would like to express my gratitude to my supervisors Hubert, Matthijs, Markus, Tanja and Hongkai for their support during my entire research and Shervan for his support during the first few months. Thanks for your endless enthusiasm, realism, humour and valuable input. Thanks for being so approachable and for all discussions we had about my thesis.

I am enormously grateful to the UTM (Universiti Teknologi Malaysia), for letting me use their equipment during my field work in the Perak and Johor state. Thank you, Taufik and Ady for driving me through the most remote areas of Malaysia's peninsula, doing discharge measurements for five days in a row in the burning sun, feeling responsible for me because I decided to climb a mountain bare feet by myself, teaching me about the Malaysian culture and language, taking me to a traditional Malaysian wedding and letting me taste the best traditional Malaysian food. Thank you, Jacqueline and Remko for the organization, support and company during this unforgettable experience.

I want to thank my beloved fellow students and friends, of the notorious 4.84 office, the best working environment I could wish for. Thanks for the fruitful thesis related discussions, input and encouragement. But foremost thanks for the humongous amount of distraction you facilitated. Thanks for all of the non-thesis related conversations we had, your amazing stories, jokes, discussions, endless laughter, philosophical theories, epic lunches with home baked bread, fresh salmon, hummus, mozzarella, cream cheese avocado bagels and so much more. Thanks Emma, Laurène, Stephanie, Tim, Anna, Sadie, Wouter, Marianne, Nadine, Maarten, Bouke, Marij, Gert and Dirk.

Thanks to all of my roommates who made it so nice to come home: thanks Kwido, Rendier, Circus, Zwakkie, Brooz, Reiger, Boes, Priveetje and Els. Thank you dear members of A05 for the often fruitful coffee breaks: thanks Ali, Bart, Daan, Jorn. Dear clubgenoten, thank you very much for motivating me throughout this research, thanks Loes, Daan, Bir, Eef, Anne, Els, Elien, Door, Cath, Anouk, Kel, and Nyn. And thank you Michiel for the extra distraction during the crucial final months in Amsterdam.

Last but certainly not least, I am grateful for the infinite support of my parents Willem and Jacqueline and my two favourite brothers, Jorrit and Gerben throughout my entire studies in Delft.

Marijn

Delft, March 2014

Table of contents

List of used Abbreviations

AC	Autocorrelation
DEM	Digital Elevation Model
ECMWF	European Centre for Medium-Range Weather Forecasts
FDC	Flow Duration Curve
FR	Fast reservoir
GPS	Global Positioning System
GR	Groundwater reservoir
HAND	Height Above Nearest Drainage
IR	Interception reservoir
JAXA	Japan Aerospace Exploration Agency
JPS	Jabatan Pengairan dan Saliran Malaysia / Department of Irrigation and Drainage Malaysia
MSL	Mean Sea Level
NASA	National Aeronautics and Space Administration
NSE	Nash-Sutcliffe Efficiency
PR	Precipitation Radar
PUB	Predictions in Ungauged Basins
SRTM	Shuttle Radar Topography Mission
STEAM	Simple Terrestrial Evaporation to Atmosphere Model
TMI	TRMM Microwave Imager
TMPA	TRMM Multisatellite Precipitation Analysis
TRMM	Tropical Rainfall Measurement Mission
UR	Unsaturated root zone reservoir
UTM	Universiti Teknologi Malaysia
VIRS	Visible and Infrared Scanner

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1. Hydrological modelling

Hydrological modelling is about understanding the hydrological cycle: what happens when a rain drop hits the ground? Conceptualising the processes that happen, translating this into mathematical equations, subsequently simulating reality and finally comparing it with observations is what hydrological modellers do. Models are used because the hydrological reality is extremely complicated. Models increase the understanding of the processes that take place and the effect of changing conditions, e.g. the effects of land-use or climatic change. Models are also of great importance to do a water resource assessment: how much water is available in the catchment and what are the flood risks? How will climatic change influence the water quantity? Finally, hydrological models are important in order to do planning and management: how will system interventions affect system performance (Beven, 2001b)?

1.2. Conceptual models

Many types of models exist: empirical, conceptual, stochastical or physically-based. In this thesis conceptual models are dealt with, models in which processes are mathematically described and where storages are modelled as reservoirs. Every reservoir has a closing water balance (Shaw, 1994). Conceptual models have the benefit of limited data demand and are therefore widely applicable in operational hydrology, also they are generally easy to understand by a non-expert (Bergström, 1992). A said disadvantage is that conceptual models demand relatively more calibration to determine the parameter values (Refsgaard *et al.*, 1989), but this is challenged by Savenije (2009b).

1.3. Challenges

All hydrological models require parameter calibration. This procedure is inevitable as models are merely an approximation of reality, due to which parameters do not directly represent measured physical properties at the scale of the schematization (Beven, 2000, 2001a, 2002a,b). Calibration involves the processing of forcing data by the model and the adjustment of parameter values in order to produce a reliable simulation. This goodness of fit is evaluated by comparing the modelled variables with a time series of the corresponding observed variables. One can imagine that the calibration process is sensitive to various uncertainty sources, e.g. uncertainty in the forcing data (such as precipitation and potential evaporation), model structure, model parameters and the observed calibration/validation data. These uncertainties directly translate into biased parameter estimates (e.g. Kavetski *et al.*, 2006a,b; Vrugt *et al.*, 2008; Thyer *et al.*, 2009) and these biased parameter estimates necessarily lead to biased model predictions.

This research focuses on the latter error source, the uncertainty in calibration/validation data. Commonly this data exists of discharge data measured at the outlet of the catchment, which is a convenient variable since it represents the integrated catchment behaviour. In this way, the lumped parameters can be calibrated at the appropriate scale (Wagener, 2007). Discharge can be highly uncertain, due to several combined sources, among others: measurement errors in stage and velocity during individual gauging, assumptions on the form of a rating curve, extrapolation of the curve and variability of the cross-section (McMillan *et al.*, 2010). Previous studies on discharge uncertainty were on the determination of uncertainties in the rating curve parameters, whereby a single rating curve definition of a specific form was used (e.g. Venetis, 1970; Petersen-Øverleir, 2004; Moyeed *et al.*, 2005; Pappenberger *et al.*, 2006; Reitan *et al.*, 2006, 2009; Di Baldassarre *et al.*, 2009; Krueger *et al.*, 2009; Liu *et al.*, 2009). Other studies investigated the possible effects of rating curve uncertainty on the calibration of the model parameters and the predictions from the calibrated models (e.g. Aronica *et al.*, 2006; Montanari, 2004).

In this research we approach this challenge differently. Instead of calibrating on the error-prone discharge data, we calibrate directly on stage. Drawback of this method is the inevitable addition of rating curve parameters, enlarging the risk of over-parameterisation (e.g. Beven, 1993, 1996, 2001a; Savenije, 2001). However, different from 'normal' conceptual model parameters, these parameters have a clear physical meaning and have been well-defined in literature. Additionally, when these parameters can unambiguously be identified, an increased risk of over-parameterisation is very low.
Calibration on stage could be a suitable approach not only for scarcely gauged basins to reduce discharge uncertainty, but also for ungauged basins, where no ground measurements are available at all. There is a great need for hydrological models in ungauged basins to assess water resources, flood and drought risk, and effects of changes in river basins due to e.g. human influences of climatic change. Many initiatives have been taken already, which have been reported in PUB (Sivapalan, 2003), a research initiative launched in 2003 by the International Association of Hydrological Sciences. In an ungauged basin radar altimeters could be used for directly measuring stage variation in (large) rivers. Subsequently, the tested models could be calibrated on these time series.

1.4. Outline

This thesis consists of three main chapters, each discussing one of the three hypotheses. Each chapter contains an introduction, methods, results, discussion and conclusions. Prior to these chapters there is a chapter on the study area and this thesis is finalized with concluding remarks.

Chapter 2 - Study area

Chapter 2 describes the study area for which the hypotheses are tested and elaborates on the data availability and reliability.

Chapter 3 - Hypothesis one: Calibrating a rainfall-runoff model on stage reduces discharge uncertainty in poorly gauged basins

Calibrating a model on uncertain discharge observations leads to biased model parameter estimates, which directly lead to biased model predictions. Chapter 3 discusses a novel method in which a model is directly calibrated on stage in order to check the following hypothesis: Calibrating a rainfall-runoff model on stage reduces discharge uncertainty in poorly gauged basins.

Chapter 4 - Hypothesis two: Topography driven rainfall-runoff models are better transferable than lumped rainfall-runoff models

It is argued whether validation on an independent period proves something, since it is very likely that the model will give satisfactory performance, as the model is calibrated on data from the same location during calibration (e.g. Pokhrel *et al.*, 2011). Therefore chapter 4 tests a different way of validating, whereby the developed models are tested in a different catchment. The hypothesis we test is the following: Topography driven rainfall-runoff models are better transferable than lumped rainfall-runoff models.

Chapter 5 - Hypothesis three: Linking a rainfall-runoff model to a steady state salt intrusion model offers mutual model improvement

Chapter 5 discusses a novel validation method, whereby a rainfall-runoff model, which predicts the fresh water flow into an estuary, is coupled to a steady state salt intrusion model, which predicts the salt intrusion length on the basis of among others, the fresh water discharge. Chapter 5 aims at highlighting the vast variety of mutual opportunities offered by the coupling of these two fields of study. The corresponding hypothesis is: Linking a rainfall-runoff model to a steady state salt intrusion model offers mutual model improvement.

Chapter 6 - Concluding remarks

Chapter 6 shortly summarizes the most important findings of this thesis.

CHAPTER 2

Study area

2.1. Introduction

The hypotheses described in Chapter 1 have been tested on the Endau River catchment, located in the State of Johor in the South-East of peninsular Malaysia (Figure 1a). The most important reason for choosing this location is the close relations with the 'Universiti Technology Malaysia' (UTM) due to which field work could easily be performed. An additional reason is the increasing flood risk in Malaysia, due to changing physical characteristics of the hydrological system caused by human influence: urbanization of the catchments and deforestation. Flood forecasting and warning systems are used, but proved to be unable to predict floods. Better models are needed to improve the ability of flood prediction (Chan, 1997). This chapter elaborates on the catchment characteristics as well as data availability and reliability.

2.2. Climate, topography and geology

Peninsular Malaysia experiences a wet and humid tropical climate that is characterized by a high annual precipitation (2000 – 4000 mm/a), a high humidity (82 – 86 %) and a rather stable high temperature (22 – 32 ˚C). The peninsula is exposed to the southwest monsoon from roughly May to September and the northeast monsoon from November to March. Areas exposed to this northeast monsoon are wetter than those exposed to the southwest monsoon (Suhaila and Jemain, 2007). The area studied here is located in the South-East and is therefore wetter than areas on the West, but drier than areas in the North-East. The catchment receives an average annual precipitation of about 2700 mm/a (based on TRMM data see Section 2.4.2.). Heaviest rainfall is experienced at the beginning of the northeast monsoon in the months November – January, while drier periods are observed during June and July. The highest rainfall recorded on one day in this catchment was 430 mm/d over a measuring period of 21 years (Suhaila and Jemain, 2007).

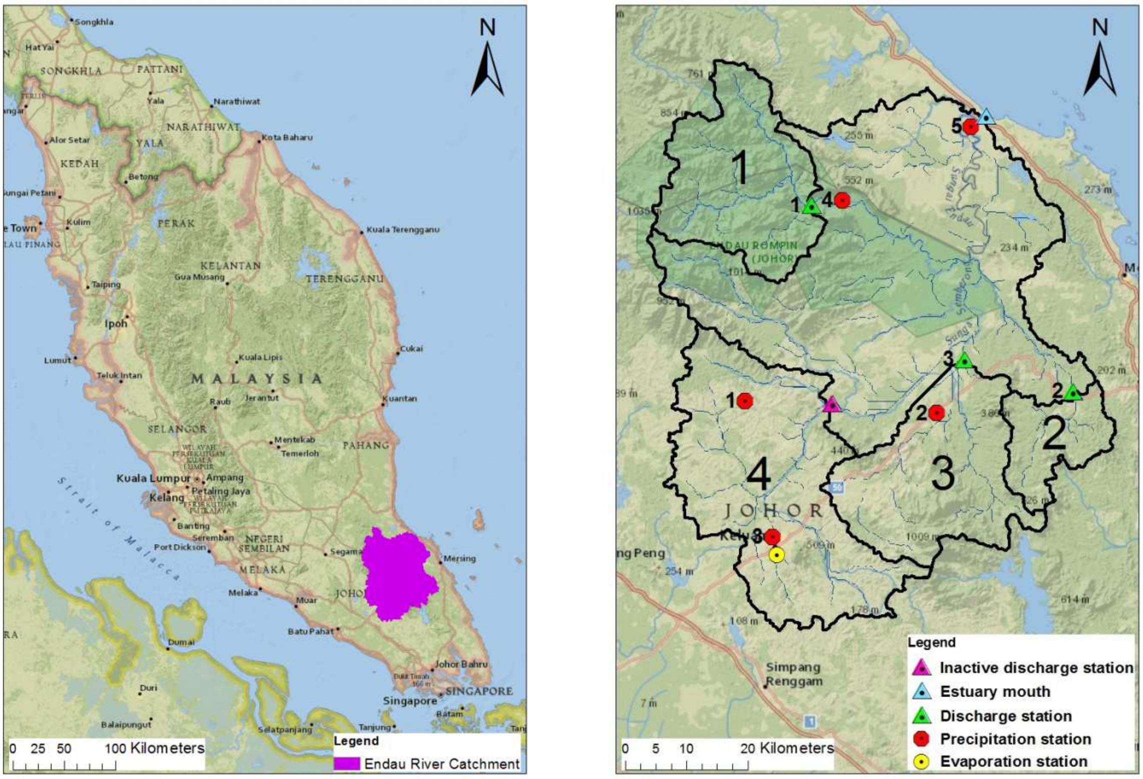
The study area boundaries are based on the area that drains into the estuary mouth through the Endau River (see Figure 1B and Appendix A for an elaboration). The full area has a size of almost 4500 km^2. The Endau River finds its origin in the mountain chain in the North-West. The Endau River has various attributes, among others the Sembrong, Lenggor and Emas River. Each of these rivers has its own sub-catchment within the full study area, indicated with numbers 1 to 4 in Figure 1b. This study focuses on sub-catchment 1 and 2, with an area of 630 km^2 and 214 km^2 respectively.

The catchment has a mountain range in the North-West and a smaller one in the South-East. The elevation ranges from 0 to 1000 m above MSL (Figure 2a). The largest part of sub-catchment 1 is covered by the Endau Rompin National Park and therefore undisturbed. The dominant vegetation on the hillslopes is tropical rainforest. On the lower lying areas grass and crop land is dominant. Sub-catchment 2 has been more disturbed as a result of human activities: large parts of the area are covered by palm oil and rubber plantations. The undisturbed parts have the same vegetation as sub-catchment 1: tropical rainforest, grass and crop land.

The catchment's soils are dominated by planosol (ISRIC – World Soil Information, 2013), a type of soil with a rather course texture that shows signs of periodic water stagnation. This type of soil usually overlies denser, finer and slowly permeable clayey sub soils (FAO, 2006).

2.3. Landscape classification

During a field visit from the 25th till the 29th of March, various locations in the catchment were classified into the three landscape classes defined by Savenije (2010), being wetland, hillslope and plateau. This in-field classification has been compared with a classification based on slope and HAND (Hight Above

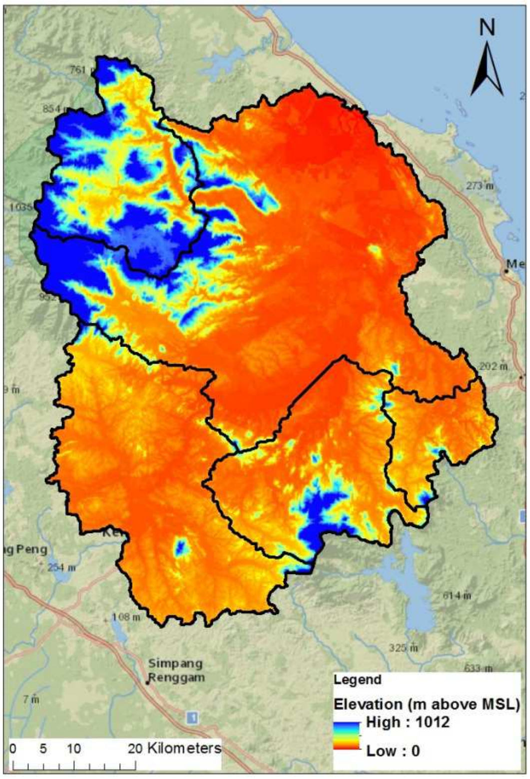


a) b)

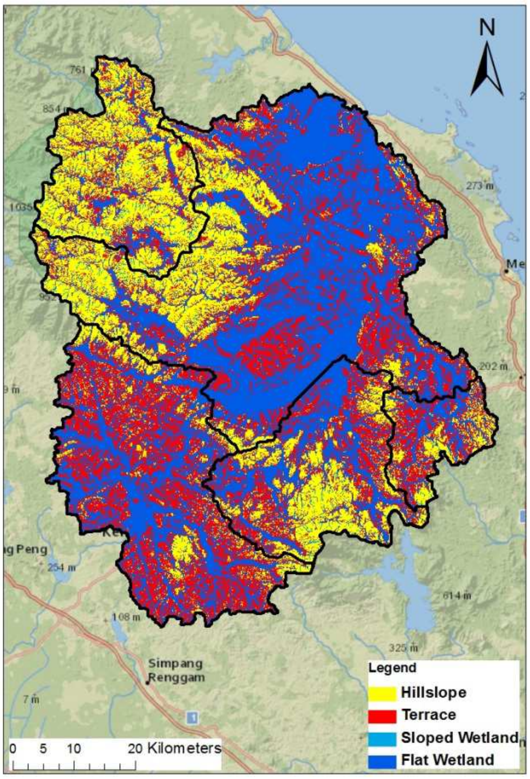
Figure 1 – a) The location of the Endau River catchment in the South-East of peninsular Malaysia and b) the location of sub-catchments and measuring stations.

Nearest Drainage) (Rennó *et al*., 2008). The slope and HAND are based on a digital elevation model (DEM) that was obtained from radar altimetry in the Shuttle Radar Topography Mission (SRTM). The DEM data from this space mission covers most of the populated regions of the world and is freely available at a spatial resolution of 3x3 arc seconds, equal to about 90x90 meters around the equator. The DEM data has a vertical accuracy of 16 meters and a horizontal accuracy of 20 meters (Berry *et al*., 2007). The stream initiation threshold required for the HAND algorithm has been set at 15 upstream cells, equal to 12 hectares. The thresholds for HAND and slope have been set at 5.9 m and 0.129, respectively, similar to what Gharari (2011) did. An analysis has been performed on the sensitivity of these thresholds, from which can be concluded that the sensitivity is low. This low sensitivity implies that there is a clear distinction in between the classes. Errors caused by a wrong set of thresholds will be negligible compared to other errors induced.

Sub-catchment 1 is classified into 59% hillslope, 20% terrace and 21% wetland and sub-catchment 2 into 27% hillslope, 38% terrace and 35% wetland, on the basis of threshold values for HAND and slope, as shown in Figure 2b and Table 1. In Table 1 the wetland is again subdivided in sloped and flat wetland, similar to Gharari (2011), however, in this catchment the percentage of sloped wetland is very limited. The classification made during the field campaign corresponds with the classification based on the threshold values.



a)



b)

Figure 2 – a) The elevation and b) the classification map of the Endau River catchment.

Table 1 – Classification of the full catchment and its sub-catchments into hillslope, terrace, sloped wetland and flat wetland.

Class	Full catchment	Sub catchment 1	Sub catchment 2	Sub catchment 3	Sub catchment 4
Hillslope	25%	59%	27%	30%	11%
Terrace	30%	20%	38%	30%	46%
Wetland	45%	21%	35%	40%	43%
Sloped wetland	1%	4%	1%	2%	0%
Flat wetland	44%	17%	34%	38%	43%

2.4. Data availability and reliability

The Endau River catchment is considered a poorly gauged basin, due to a relatively scarce amount of reliable hydrological data. The catchment has three ground stations for stage/discharge, five for precipitation and one for evaporation, of which the locations are indicated in Figure 1b. Satellites and climatic models provide additional meteorological data series for the Endau River catchment. An overview of the data characteristics is given in Table 2, where the last column indicates which data has actually been used in this research.

Table 2 – Overview of the available hydrological ground data.

Station	Observed variable	Data source	Elevation (m)	Observation interval	Start	End	Missing data (over 2003-2013)	Used
1	Discharge	JPS (ground data)	36	Daily	1961	2013	27%	
2	Discharge	JPS (ground data)	34	Daily	2000	2013	18%	
3	Discharge	JPS (ground data)	20	Daily	1978	2013	24%	
1	Stage	JPS (ground data)	36	Daily	1961	2013	27%	√
2	Stage	JPS (ground data)	34	Daily	2000	2013	18%	√
3	Stage	JPS (ground data)	20	Daily	1978	2013	24%	
1	Evaporation	JPS (ground data)	37	Daily	1962	2013	32%	
-	Evaporation	STEAM (model based)	-	Daily	2003	2013	0%	√
1	Precipitation	JPS (ground data)	63	Daily	1980	2013	10%	
2	Precipitation	JPS (ground data)	29	Daily	2003	2013	24%	
3	Precipitation	JPS (ground data)	31	Daily	2000	2013	9%	
4	Precipitation	JPS (ground data)	36	Daily	1975	2013	12%	√
5	Precipitation	JPS (ground data)	11	Daily	1970	2013	16%	
-	Precipitation	TRMM 3B42 (satellite data)	-	3 hourly	2003	2013	0%	√

2.4.1. Discharge, Stage

Average daily stage is observed at four stations, all operated by the Department of Irrigation and Drainage Malaysia (JPS), of which three are still active. During a field survey, various JPS stations were visited all over peninsula Malaysia of which two in the Endau River catchment. All visited JPS stations measure at unlined locations. Various stations measure in the vicinity of a bridge, whereby some stations are located downstream and others upstream of this bridge, risking the effect of backwater to be included in the measurements. On most locations both an automatic sensor and a manual staff gauge are present, it is assumed that the automatic logger is commonly used, as no observers were spotted during the surveys. In the Endau River catchment station 2 was the only well accessible gauging station, here we performed additional discharge measurements and surveyed the cross-section, see Appendix B.

This study focuses on station 1 and 2, with corresponding sub-catchments 1 and 2. At station 1, stage has been measured since the year 2000 and at station 2 since the year 1961. The stage data from station 1 and 2 have 27% and 18% of missing data, respectively. These gaps were not filled in this research, but were disregarded. JPS converts the stage readings into discharge using a stage-discharge relation (a rating curve). The rating curves used over time are unknown to the authors. However, when plotting the observed stage versus the observed discharge, one can distinguish the various curves that were used over time. Visualisation of the time series together with the rating curves provide better understanding of the reliability of the data.

Figure 3 and Figure 4 show the time series of discharge and stage for sub-catchment 1 and 2, respectively. At first sight, the discharge and stage data of sub-catchment 1 look reliable until 2007, while after that year a lot of data is missing and the behaviour changes. A closer look reveals that the peaks could be too high, caused by a wrong extrapolation of the rating curve. The stage data of sub-catchment 2 looks rather stable and reliable. The maximum observed value is about 17 m, which is about 5 m higher than the observed water level during a field campaign. From site inspection including investigation of erosion patterns and preventive measures against erosion, it can be concluded that a 5 meter higher water depth is not unlikely to occur. The corresponding discharge data for sub-catchment 2 look on the other hand very

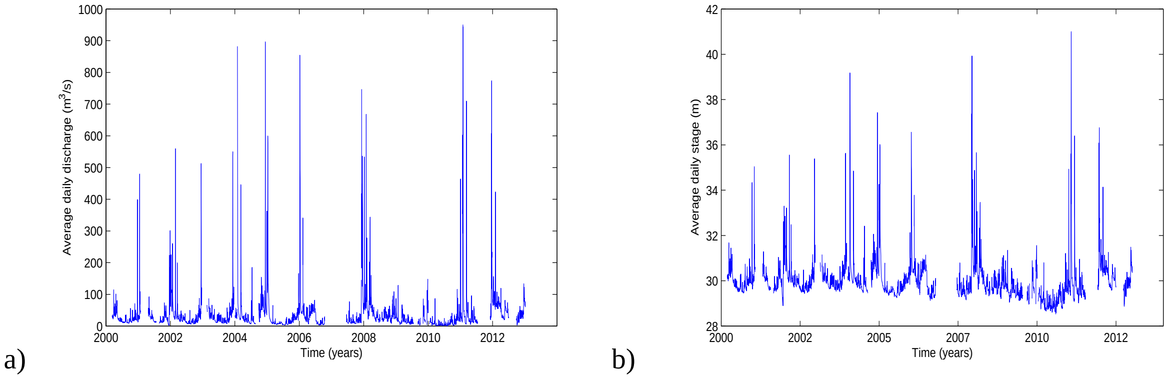


Figure 3 – a) Average daily discharge and b) average daily stage of sub-catchment 1 over the entire gauging period.

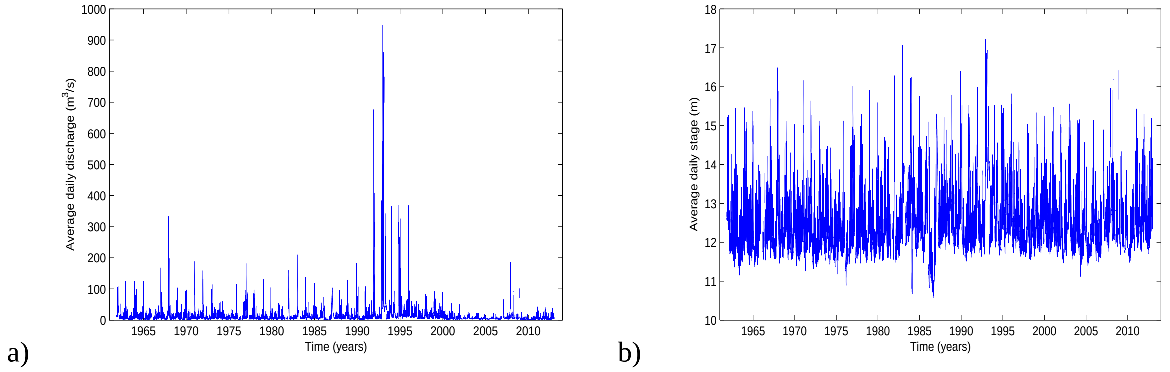


Figure 4 – a) Average daily discharge and b) average daily stage of sub-catchment 2 over the entire gauging period.

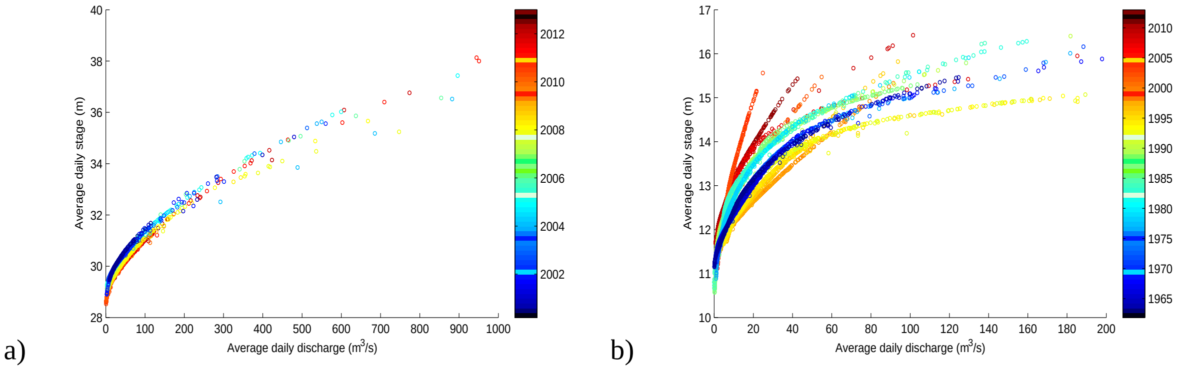


Figure 5 – a) The rating curves used at discharge station 1 over the period 2000 – 2013 and b) rating curves used at discharge station 2 over the period 1961 – 2013. The different colours indicate in which year a certain rating curve was used.

messy. It can be seen that the peak discharges at sub-catchment 2 fluctuate tremendously over the years; this is far from likely behaviour. Figure 5 shows the stage plotted versus the discharge for both sub-catchments, whereby the different colours indicate the year in which the rating curve was used. From these graphs it becomes clear that the rating curves for sub-catchment 1 have been rather stable over the years, while at sub-catchment 2 the variability, mainly in the high flows is unrealistically high. In 1995 a stage reading of 14 m gives a discharge of 50 m^3/s, while 5 years later the same stage reading gives a discharge of only 17 m^3/s. In order to realize this change in discharge, the cross-sectional area must have changed tremendously within only 5 years, this is very unlikely. Therefore the sub-catchment 2 rating curves are treated as far from reliable; since the discharge is calculated from stage using these rating curves, the discharge of sub-catchment 2 is considered unreliable as well.

2.4.2. Precipitation

Precipitation is available from two sources: ground and satellite data. For the satellite data the product 3B42 is used, which has been obtained in the Tropical Rainfall Measurement Mission, a joint space mission of the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). Ever since its launch in 1997, TRMM is providing rain radar and microwave radiometric data that gives the vertical distribution of precipitation over the tropics in a band between 25 degrees north and south of the equator (NASA, 2012). The relevant instruments that are used for precipitation estimations are the Precipitation Radar (PR), the TRMM Microwave Imager (TMI) and the Visible and Infrared Scanner (VIRS). PR yields information on the intensity and distribution of rain, TMI quantifies water vapour, cloud water and rainfall intensity in the atmosphere and VIRS measures radiation coming up from the earth (from visible to infrared) which serves as an indicator of rainfall. TRMM based precipitation estimates are created using the TRMM Multisatellite Precipitation Analysis (TMPA), which has led to precipitation estimates at a minimum temporal scale of 3 hours and a minimum spatial scale of 0.25° x 0.25°. In this study, 3-hourly precipitation estimates have been used that were corrected by ground data (3B-42 V7.1). The Endau River catchment fits within a square of 4 by 4 TRMM cells.

The ground data is measured with a tipping bucket at five locations, as indicated in Figure 1b. During a field campaign various rain stations were visited, from which can be concluded that not all rain stations are set up according to the standards. Often the stations do not meet the rule of thumb indicating that no obstacles should be present nearer than about five times the obstacle height, as this will influence the air flow and therefore the rainfall in the gauge (Luxemburg *et al.*, 2011), see the illustrative photo in Appendix B. Missing data (9 – 24%) in the ground series has been filled up by taking the value measured by the nearest precipitation station, if that station also had a missing value, the next closest station was taken and so on. Not all the precipitation gauages measure the same amount of rain. In the filling up process this has been accounted for by scaling the data accordingly. For the days at which all gauging stations had a missing value, the missing value was set to 0. Missing data in the TRMM series were barely found; the few missing values have been replaced with 0.

The precipitation measurements have been checked for irregularities by means of a double mass analysis, whereby the accumulated values of the station under investigation have been plotted against accumulated values of the average of other stations over the same period. In the ideal situation the double mass curve should follow a straight line, which means that the investigated station is affected to the same extent by the same trends as the other stations. A break in the slope of the curve would indicate that conditions at the

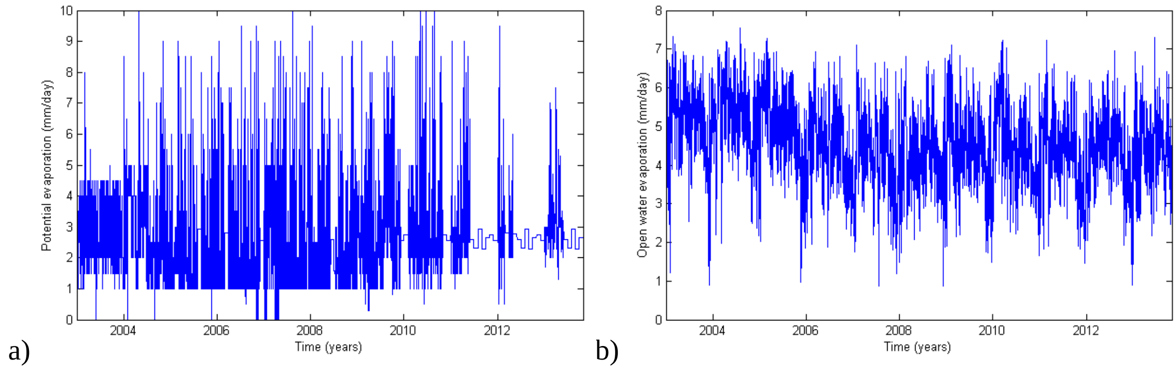


Figure 6 – Potential evaporation based on a) pan evaporation and b) Penman open water evaporation.

Table 3- The long term annual average precipitation measured by gauge and satellite.

Station	Long term annual average precipitation (mm/a)	
	Gauge	TRMM
1	1875	2782
2	1932	2820
3	1661	2782
4	2953	2763
5	2506	2649

station under investigation changed, while at other stations this was not the case. Appendix C shows the double mass plots of both the satellite and the filled up ground data. The double mass curves of the satellite data form almost perfect straight lines, while the gauge data show more irregularities. Least irregularities are observed at gauge 4. The long term annual averages of the gauge and the satellite data are shown in Table 3. Note that the long term average of the gauge is compared with the long term average of the geographical TRMM cell in which the gauge is located. We see that gauge 4 and 5 show a rather good resemblance, while gauges 1, 2 and 3 measure on average way less than the satellite does. In Section 2.2. it became clear that the long term annual average precipitation on peninsula Malaysia ranges from 2000 to 4000 mm/a, whereby the East coast is classified as the wettest zone due to the northeast monsoon. Since the Endau River catchment is located on the East coast, it can be concluded that the precipitation measured by gauges 1, 2 and 3 is too low. Due to the deviation in the long term annual average and the bad performance on the double mass analyses, it is concluded to classify the data series from rain gauges 1, 2 and 3 as unreliable.

2.4.3. Evaporation

Potential evaporation is available from two sources: pan evaporation and open water evaporation calculated with the Penman equation (Penman, 1948). Both time series are shown in Figure 6 over the period 2003 to 2013. The first series has been measured ever since the year 1962 at a location indicated in Figure 1b, however, 32% of the measurements is missing. These gaps are filled by the average of that very month. When more than half of the data of that month is missing, the long term monthly average is used instead. Potential evaporation based on Penman is calculated in geographical cells at 1.5 degree spatial resolution using STEAM (Simple Terrestrial Evaporation to Atmosphere Model) (Wang-Erlandsson *et al.*, 2014). The meteorological data (radiation, temperature, relative humidity and wind speed data) required as model forcing is obtained through ERA Interim (Dee *et al.*, 2011), the latest global atmospheric reanalysis produced by the ECMWF. Berrisford *et al.* (2009) provided a detailed description of the ERA-Interim product archive; information on the current products and data avialability can be found at: http://www.ecmwf.int/research/era. In order to make the open water evaporation suitable to use as potential evaporation, the data series are corrected by a factor 0.8, according literature values (Mohamed *et al.*, 2012). The Penman open water evaporation is available from 2003 until present.

We observe large differences between the two time series in absolute magnitude and in annual cycle. In fact, no clear cycle was observed in the pan potential evaporation, while the Penman potential evaporation shows a clear annual variation: lower potential evaporation values during the monsoons, when temperatures are lower and higher values out of the monsoon season, when higher temperatures are experienced. The average annual potential evaporation is roughly 950 mm/a according to the potential evaporation based on pan evaporation and roughly 1450 according to Penman. When comparing the typical daily potential evaporation with literature values, the pan evaporation is too low, while the Penman potential evaporation gives a good estimate. Additionally, when observing the pan evaporation, we see that a large amount of measurements has been rounded to the value of 1 mm/d. This could have different explanations, which all lead to one conclusion: these measurements look far from reliable. It is therefore decided to use the Penman potential evaporation in this study.

2.5. Conclusion on data usage

Based on the data reliability, it is decided to reject the uncertain discharge data and use the stage data for calibration. For the potential evaporation the Penman open water evaporation is used, corrected by a factor 0.8 (Mohamed *et al.*, 2012). The precipitation input should preferably originate from ground data, as this is considered the truth. Gauge 4, located in the vicinity of sub-catchment 1, shows good correlation with the TRMM data and shows minor irregularities on a double mass curve, therefore, the time series of this gauge are used for sub-catchment 1. Since the gauging stations in the vicinity of sub-catchment 2 are considered unreliable, it is decided to use TRMM data instead. The downside of using TRMM data is that the precipitation has been averaged over the entire area, as a result of which peaks are flattened out. It was therefore decided to test the gauged data as well, despite the unreliability of the rain gauges closely located to sub-catchment 2. Using these precipitation series, the models gave very bad performance and hence the data from the rain gauges was rejected.

CHAPTER 3

Hypothesis one: Calibrating a rainfall-runoff model on stage reduces discharge uncertainty in poorly gauged basins

3.1. Introduction

Direct measurement of river discharge is time consuming and financially demanding. Continuous river discharge measurements are therefore generally derived from continuous stage measurements, through a stage-discharge relation, also called a rating curve. Rating curves are determined by fitting a curve to a limited number of points (h_i, Q_i), whereby h_i and Q_i represent stage and discharge, measured in a certain cross-section of a river at a fixed geographical location. Commonly these points (h_i, Q_i) originate from measurements done under regular flow conditions, due to which a considerable part of the curve is based on interpolation and extrapolation (e.g. Leonard *et al.*, 2000; Pappenberger *et al.*, 2006; Shrestha *et al.*, 2007). Therefore the uncertainty in the rating curve particularly during floods can be considerable, which directly translates into uncertainty in the flow data.

In hydrological modelling a hydrograph (discharge versus time) is commonly modelled using forcing data (e.g. precipitation and potential evaporation), together with a model structure and model parameters (Figure 7a). Subsequently, this modelled hydrograph is compared with an observed hydrograph to check its performance. However, as we just discussed, discharge is not directly measured, but originates from stage readings. A rating curve model is required to translate stage into discharge, schematized in Figure 7b. Uncertainty in this rating curve model directly results in uncertainty in the 'observed' hydrograph. Calibrating a model on uncertain discharge observations leads to biased model parameter estimates, which directly lead to biased model predictions.

This chapter shows an approach whereby a conceptual rainfall-runoff model is calibrated on the basis of stage data only. In addition to the model parameters, extra parameters are added that define the rating curve. Subsequently, the modelled hydrograph is translated into modelled stage through this 'reanalysed' rating curve, this conceptualization is schematized in Figure 7c. Finally, the modelled stage is compared with the observed stage. The above explanation is mainly used to provide clarity, for practical reasons it was decided to work slightly differently, although the outcome remains the same: once the rating curve parameters have been calibrated, the 'reanalysed' hydrograph is established using the observed stage readings. Subsequently, the model hypotheses on catchment behaviour are tested by conventional methods.

3.2. Methods

3.2.1. Model descriptions

To check the influence of the model structure on the process of stage calibration, two conceptual models have been tested for sub-catchment 1 and 2. A topography driven semi-distributed model has been used that distinguishes between different topographical classes and a lumped model that treats the catchment as a whole. The structure of the topography driven model is similar to the FLEX^T model structure developed by Gao (2013) and was already successfully applied to catchments in North-West China and Thailand. The lumped model is similar to the HBV concept (Bergström, 1992). Both models operate on daily timescale. Although timing may be important, no lag function has been included in any of the models, as the scarcity of the rainfall data does not allow to estimate the timing. This issue was implicitly validated by calibration on the flow duration curves (Section 3.2.5.1.).

The upcoming sections describe the two conceptual models. Since there are a lot of similarities between the lumped and the topography driven model, it was decided to extensively elaborate on the lumped model description, and only explain the new or different conceptualisations in the description of the topography driven model.

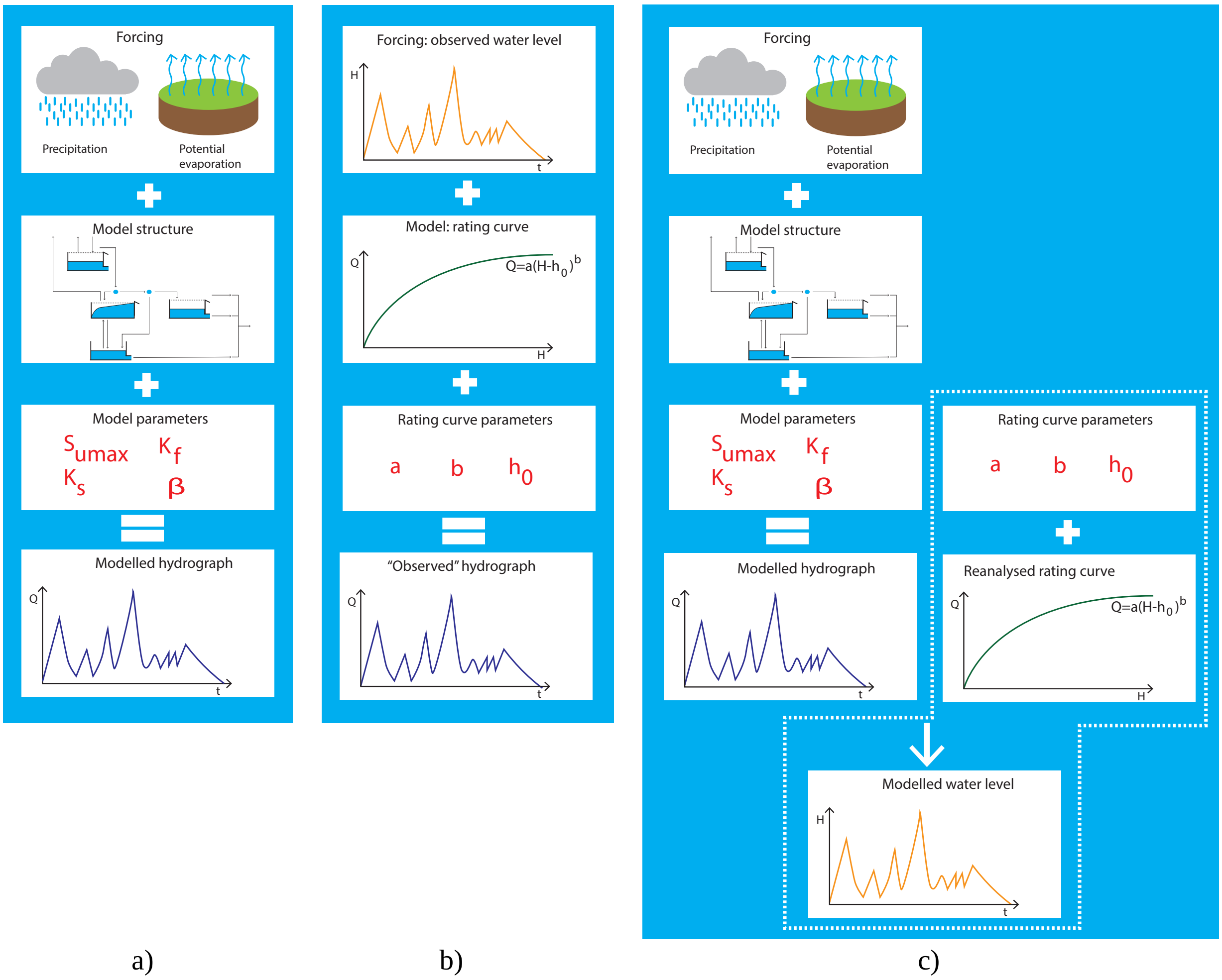


Figure 7 – a) Schematised conceptualisation of the modelling process, whereby a modelled hydrograph is formed. b) The construction of the 'observed' hydrograph from water level measurements using a rating curve model. c) The novel approach whereby water level is modelled rather than a hydrograph.

3.2.1.1. Lumped model

For the lumped model, the best results have been obtained with the four reservoir model of which a schematic representation is depicted in Figure 8. This lumped model comprises four reservoirs: an interception reservoir (IR), an unsaturated root zone reservoir (UR), a fast reservoir (FR) and a groundwater reservoir (GR).

Interception module

The interception module was conceptualised as a threshold process as proposed by Savenije (2004). Precipitation first reaches IR, of which the water balance is given in Eq. 1. IR has a maximum storage capacity represented by I_{max}. When the interception storage exceeds I_{max}, the reservoir spills and the excess precipitation takes part in subsequent processes, such as infiltration and surface runoff. The intercepted water can evaporate from IR as long as there is water and energy available, according Eq. 2.

(1) $$\frac{dS_i}{dt} = P - E_i - P_e$$

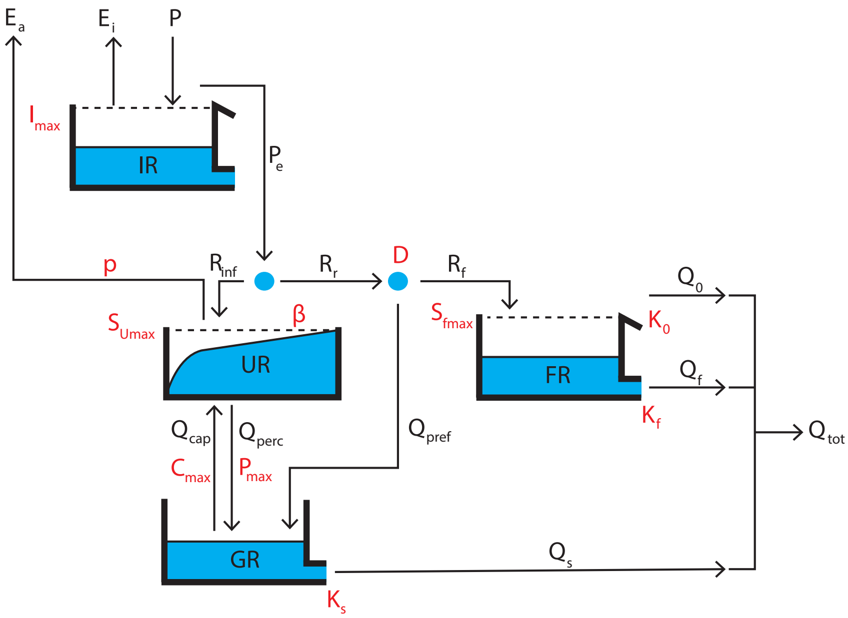


Figure 8 - Lumped model structure including the model parameters (red) and model fluxes (black).

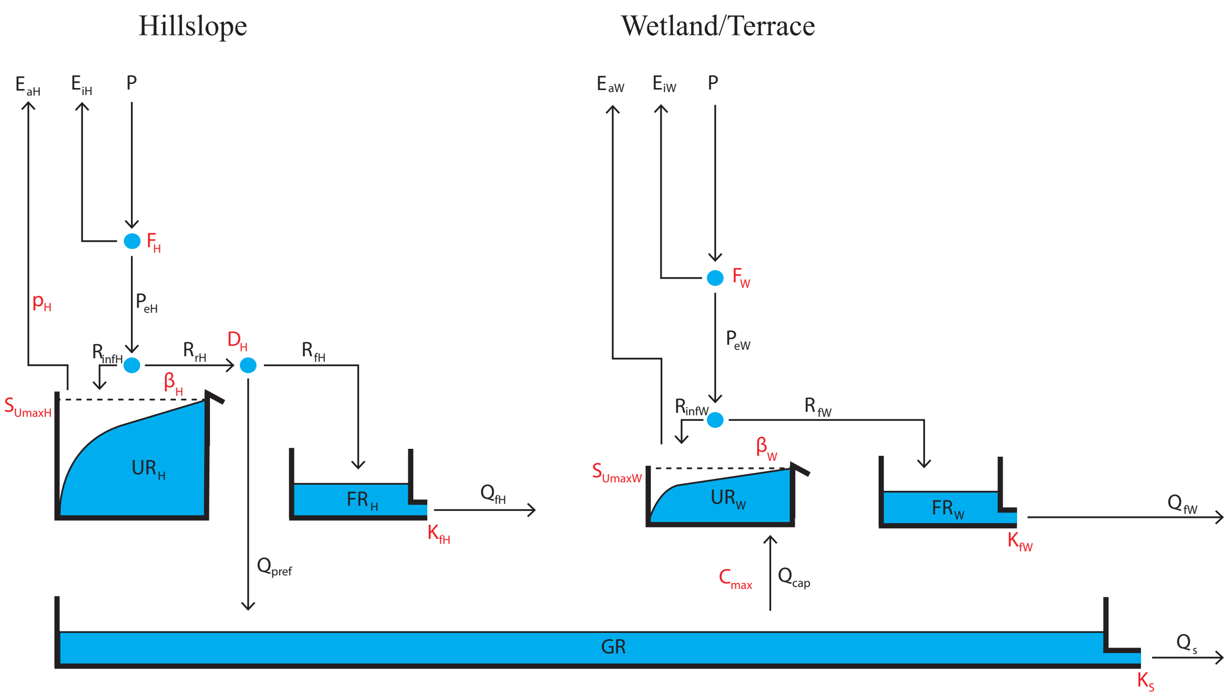


Figure 9 - Topography driven model structure including the model parameters (red) and model fluxes (black).

(2) $E_i = \min\left(\frac{S_i}{\Delta t}, E_p\right)$

(3) $P_e = \max\left(S_i - I_{max}, 0\right)$

In which $\frac{dS_i}{dt}$ (mm/d) is the storage change in IR, E_i (mm/d) is the interception evaporation, Δt (d) the calculation time step, S_i (mm) is the storage in IR, E_p (mm/d) is the potential evaporation, P_e (mm/d) is the effective precipitation and I_{max} (mm) is the maximum storage capacity of IR.

Unsaturated root zone module

When the interception storage exceeds the threshold I_{max}, there is effective precipitation, P_e, that enters UR or becomes runoff depending on the saturation state of UR. The fraction (ρ) of effective precipitation that comes to runoff (R_r) is determined through a nonlinear function of the relative root zone moisture content ($S_u/S_{u,max}$) of the unsaturated reservoir and a shape parameter β. When UR is filled to capacity, all excess rainfall is routed to FR or GR. The part of the effective rainfall that does not infiltrate is partitioned into a fast component into FR (R_f) and a preferential recharge component into GR (Q_{pref}), according to the partitioning coefficient D. The actual evaporation from UR (representing soil evaporation, transpiration and possibly open water evaporation) depends on the relative root zone moisture content ($S_u/S_{u,max}$) when this ratio is below the threshold p. When a higher ratio is obtained, the actual evaporation equals the potential evaporation reduced by the energy used for interception evaporation.
Percolation from UR to GR is linearly related to the relative root zone moisture content ($S_u/S_{u,max}$) with a maximum percolation rate (P_{max}). Capillary rise from GR to UR takes place in case there is excess potential energy after interception and evaporation and is linearly related to the relative excess potential evaporation ($\frac{E_p - E_i - E_a}{E_p}$) and the maximum capillary rise (C_{max}).

(4) $\frac{dS_u}{dt} = R_{inf} + Q_{cap} - E_a - Q_{perc}$

(5) $\rho = \min\left(1, \left(\frac{S_u}{S_{u,max}}\right)^{\beta}\right)$

(6) $R_{inf} = (1 - \rho)P_e$

(7) $R_r = \rho P_e$

(8) $R_f = D\rho P_e$

(9) $Q_{pref} = (1 - D)\rho P_e$

(10) $E_a = (E_p - E_i)\min\left(1, \frac{S_u}{pS_{u,max}}\right)$

(11) $Q_{perc} = \min\left(\frac{S_u}{\Delta t}, P_{max}\frac{S_u}{S_{u,max}}\right)$

(12) $Q_{cap} = \min\left(\frac{S_g}{\Delta t}, C_{max}\frac{E_p - E_i - E_a}{E_p}\right)$

Where $\frac{dS_u}{dt}$ (mm/d) is the storage change in UR, R_{inf} (mm/d) is the infiltration into UR, Q_{cap} (mm/d) is the capillary rise flux from GR to UR, E_a (mm/d) the actual evaporation, Q_{perc} (mm/d) is the percolation flux from UR to GR, ρ (-) represents the runoff coefficient, S_u (mm) the storage in UR, $S_{u,max}$ (mm) the maximum storage in UR and β (-) is a measure to describe the spatial heterogeneity distribution in the catchment, R_r (mm/d) is the fraction of effective precipitation that comes to runoff, D (-) is a splitter that separates fast runoff from preferential flow, R_f (mm/d) is the fast runoff into FR, Q_{pref} (mm/d) is the preferential flow into GR, p (-) the fraction of $S_{u,max}$ above which the actual evaporation is equal to the potential evaporation, when the fraction is lower than p the E_p is constrained by the water available in S_u, P_{max} (mm/d) the maximum percolation rate from UR to GR, S_g (mm) the storage in GR and C_{max} (mm/d) is the maximum capillary rise from GR to UR.

Fast and slow reservoir module

The FR and GR are conceptualised as linear reservoirs to represent surface runoff (Eq. 15), fast runoff (Eq. 16) and base flow (Eq. 17), together the total flow (Eq. 18).

(13) $$\frac{dS_f}{dt} = R_f - Q_0 - Q_f$$

(14) $$\frac{dS_g}{dt} = Q_{perc} + Q_{pref} - Q_{cap} - Q_s$$

(15) $$Q_0 = \frac{1}{K_0} \max\,(0, S_f - S_{f,max})$$

(16) $$Q_f = \frac{1}{K_f} S_f$$

(17) $$Q_s = \frac{1}{K_s} S_g$$

(18) $$Q_{tot} = Q_s + Q_f + Q_0$$

Where $\frac{dS_f}{dt}$ (mm/d) is the storage change in FR, $\frac{dS_g}{dt}$ (mm/d) is the storage change in GR, Q_0 (mm/d) is the rapid surface overland flow with timescale K_0 (d), active when the storage exceeds the threshold value $S_{f,max}$ (mm), Q_f (mm/d) is the subsurface drainage of UR with timescale K_f (d), Q_s (mm/d) is the base flow from GR with timescale K_s (d) and Q_{tot} (mm/d) is the sum of the three individual components.

3.2.1.2. Topography driven model

A schematic representation of the topography driven model is depicted in Figure 9. This model is a simplification of the model used by Gao (2013) in western China and was developed on the basis of the model structures as presented by Savenije (2010). Savenije (2010) distinguished between three parallel classes: wetland, terrace and hillslope. Each of the classes was conceptualised in a model structure as such to fulfil its dominant mechanism. Since terraces were not found on the geologically young peninsula Malaysia and to simplify the conceptualisation to decrease the amount of parameters, wetland and terrace are in this research combined in one structure. For convenience we shall now refer to this combined structure as 'wetland'. The wetland and the hillslope structure run parallel but are connected through the groundwater reservoir (GR). Next to GR, both the hillslope and the wetland structure have an unsaturated root zone reservoir (UR) and a fast reservoir (FR).

Wetlands are relatively flat areas, with shallow soils, a shallow water table depth and therefore a limited residual storage capacity. The dominant runoff generation mechanism is saturation excess overland flow. This is a lateral and fast process, therefore wetlands show a fast response to precipitation. Hillslopes are mostly forested. Two important life-support functions for a forest ecosystem are drainage to maintain an aerated soil and moisture retention to bridge dry spells. To fulfil these contradicting functions, the dominant mechanism is subsurface drainage through preferential pathways. This mechanism does not cause excessive erosion, drains excess water, but enables the wetting of stagnant pockets in the soil from which the roots can tap their water (Savenije, 2010).

Interception module

Interception in the topography driven model structure is considered a threshold process, calculated as the minimum of the precipitation, the maximum interception flux, F, and the potential evaporation, all expressed in mm/d (de Groen *et al.*, 2006). The maximum interception flux has different values for the hillslope and wetland structure, F_H and F_W respectively (Eq. 19 and Eq. 20). When the precipitation exceeds the maximum interception flux, water is routed towards the unsaturated reservoir as effective precipitation ($P_{e,W}$ for the wetland and $P_{e,H}$ for the hillslope).

(19) $$E_{i,H} = \min\,(E_p, F_H, P)$$

(20) $$E_{i,W} = \min\,(E_p, F_W, P)$$

In which, $E_{i,H}$ (mm/d) and $E_{i,W}$ (mm/d) are the interception evaporation from hillslope and wetland, E_p (mm/d) is the potential evaporation, F_H (mm/d) and F_W (mm/d) are the maximum interception fluxes for hillslope and wetland and P (mm/d) is the precipitation.

Unsaturated root zone module

The runoff is conceptualised in the same way as in the lumped model. The thresholds for maximum storage in UR and the value for the coefficient β are again landscape dependent. Different is the threshold mechanism, which is added to UR. When the storage in UR exceeds the maximum storage capacity, the overflow becomes active and routes the excess water into the fast reservoir.

The hillslope structure has a splitter function that partitions the runoff in groundwater recharge and a fast component. The splitter, D_H, works the same as in the lumped model. The wetland structure does not have this splitter, since it is assumed that wetlands are seepage areas rather than infiltration areas. All the effective rainfall that becomes runoff directly routes into the fast reservoir.

The actual evaporation (representing soil evaporation, transpiration and possibly open water evaporation) in the hillslope structure is conceptualised the same as in the lumped model, however, the p factor is here called p_H, see Eq. 21. The actual evaporation in the wetland structure is equal to the potential evaporation, unless the storage in UR is not sufficient, according Eq. 22. Capillary rise only occurs in the wetland structure from GR to UR and is conceptualised as in Eq. 23.

(21) $$E_{a,H} = (E_p - E_{i,H})\min\,(1, \frac{S_{u,H}}{S_{u,max,H}p_H})$$

(22) $$E_{a,W} = \min\,(E_p - E_{i,W}, \frac{S_{u,W}}{\Delta t})$$

$$(23) \quad Q_{cap} = \min(C_{max}W, \frac{S_s}{\Delta t})$$

In which $E_{a,H}$ (mm/d) and $E_{a,W}$ (mm/d) are the actual evaporation from hillslope and wetland, $S_{u,H}$ (mm) and $S_{u,W}$ (mm) are the storages in UR of the hillslope and wetland, $S_{u,max,H}$ (mm) is the maximum storage in UR, p_H (-) is the fraction of $S_{u,max,H}$ above which the actual evaporation is equal to the potential evaporation, when the fraction is lower than p_H, the $E_{p,H}$ is constrained by the water available in $S_{u,H}$, Q_{cap} (mm/d) is the capillary rise, C_{max} (mm/d) is the maximum daily capillary rise, W (-) the part of the catchment classified as wetland and S_s (mm) is the storage in GR.

Fast and slow reservoir module

As mentioned before, the two parallel structures are connected through one ground water reservoir, again conceptualised as a linear reservoir, similar as was done in the lumped model. The outflow of the fast reservoirs in both the wetland and the hillslope structure is also conceptualised similar to the lumped model, however no threshold function is present and the time scales are again structure dependent.

3.2.2. Model parameters

Both the topography based and the lumped model are conceptual and require a set of parameters that corresponds with physiographic properties of the basin in order to generate reliable simulations (Shamir, 2005). These parameters are catchment specific and can either be determined by direct measurement in the field (those parameters that reflect measurable catchment properties, which are in practice barely present), derived from analysis of measured variables (parameters that can be directly estimated through an analysis of the input-output behaviour of the system), derived from literature or determined by automatic or manual calibration. The more parameters, the easier parameters can compensate for each other and in this way cause equifinality (over-parameterisation) (e.g. Beven, 1993, 1996, 2001a and Savenije, 2001).

3.2.2.1. Parameter ranges

The lumped model has 11 parameters and the topographical model has a total of 14 parameters to be determined. In the topography driven model the parameters indicating the part of the catchment classified as wetland (*W*) and hillslope (*H*), can directly be determined from the landscape classification, described in Section 2.3. (Figure 2 and Table 1). To limit the degrees of freedom even more and therefore decrease the chance of over-parameterisation, the *p* coefficient (the fraction of $S_{u,max}$ above which the actual evaporation is equal to the potential evaporation) of the lumped model and the similar p_H coefficient of the topographical model have been given a value of 0.5 (Savenije, 1997). The depletion of the slow reservoir, indicated by variable K_s, has been fixed using a recession slope analysis, described in Appendix D. In the lumped model, the timescale, K_0, for the surface runoff has been set at 1 day, for it was assumed that overland flow is such a fast process that it reaches the stream in (less than) a day (e.g. Savenije, 2009a). This results in 8 free parameters for the lumped model and 10 free parameters for the topographical model. The prior parameter range of both conceptual models is given in Appendix E. Note that the parameters H and W have been left out and the value for K_s has not yet been fixed in this table, as this value differs per sub-catchment.

The topography driven model has thus 2 more free parameters than the lumped model. One may therefore conclude that the topographical model will always be better capable of simulating a certain time series (Akaike, 1973; Schoups *et al.*, 2008). However, imposed parameter constraints (Section 3.2.2.2.) reduce the degrees of freedom substantially. Additionally, this study does not directly focus on comparing different models, therefore the issue of an unequal number of parameters is considered of minor importance.

3.2.2.2. Parameter constraints

For the topographical model, parameter constraints were introduced and applied after the random sampling process (Section 3.2.5.). Parameters sets that did not meet these constraints were treated as non-behavioural and rejected directly after the sampling process. These parameter constraints are based upon our perceptual understanding of the catchments, investigation of the hydrographs and expert knowledge. The interception threshold for hillslope was assumed higher than that of the wetland, as forests (assumed the most important land cover on hillslope) have a larger interception capacity than grass or crop land (assumed the most important land cover on wetland). These differences in land cover also result in the constraint whereby the maximum root zone storage of the forests ($S_{u,max,H}$) was assumed larger than that of the wetland, since forest roots deeper than grassland. Finally, there are constraints on the time scales. The time scale of the outflow from the groundwater reservoir was assumed longer than those for $K_{f,H}$ and $K_{f,W}$, since release from the groundwater reservoir is the slowest process of all. For the constraints between $K_{f,H}$ and $K_{f,W}$ various options have been tested. The best results were obtained with the constraint whereby the time scale of the hillslope was shorter than the time scale of the wetland. This could be due to a higher gradient, causing the potential energy and thus the kinetic energy to be higher as well.

(24) $F_H > F_W$

(25) $S_{u,max,H} > S_{u,max,W}$

(26) $K_s > K_{f,H}$

(27) $K_s > K_{f,W}$

(28) $K_{f,H} < K_{f,W}$

3.2.3. Rating curve parameters

Calibration on stage as described in the introduction of this chapter, requires the addition of parameters that define the relation between discharge and stage: the rating curve. In this research it has been assumed that the rating curve parameters do not vary over the modelling period (we elaborate on this topic in Section 3.2.3.3.).

3.2.3.1. Rating curve definitions

To check the sensitivity of the rating curve to the rating curve definition, two rating curve definitions have been tested. The first definition is the widely applied relation of Eq. 29, this method requires no information on the cross-section. The second relation is the Strickler-Manning equation (e.g. Abbott *et al.*, 1986), shown in Eq. 30. This method requires cross-sectional information and was therefore only applied to sub-catchment 2, where a field inspection was done and the cross-section was surveyed. The cross-section was assumed trapezium-shaped, according Eq. 31. Table 4 gives an overview of which models are developed for which sub-catchment.

Table 4 - An overview of which models are developed for which sub-catchment.

Sub-catchment	Model structure	Rating curve definition 2 parameters	3 parameters
1	Lumped		x
	Topography driven		x
2	Lumped	x	
	Topography driven	x	
	Lumped		x
	Topography driven		x

(29) $Q = a(h - h_0)^b$

(30) $Q = A\bar{v} = \frac{AR^{2/3}s^{0.5}}{n}$

With for a trapezoidal cross-section:

(31) $A = Bd + 0.5m_1d^2 + 0.5m_2d^2$

(32) $P = B + d(\sqrt{m_1 + 1} + \sqrt{m_2 + 1})$

(33) $R = \frac{A}{P}$

(34) $d = h - h_0$

Where Q (m^3/s) is the average daily discharge, h (m) is the average daily stage, h_0 (m) is the stage reading at zero flow, a (m^{3-b}/s) is a coefficient catering for primarily width of the cross-section, roughness and slope, b (-) is the exponent depending on the shape of the cross-section, A (m^2) is the wet cross-sectional area, $\bar{v}$ (m/s) the average flow velocity, R (m) the hydraulic radius, P (m) the wetted perimeter, s (-) the energy slope, n ($s/m^{1/3}$) the Strickler-Manning roughness coefficient, B (m) is the bed width, d (m) is the water depth, m_1 (-) is the side slope of the left bank, defined as the horizontal displacement in case of 1 m vertical displacement and m_2 (-) is the side slope of the right bank.

Eq. 29 has three unknown parameters which require calibration: a, b and h_0. In the second definition (Eq. 30) the Strickler-Manning roughness has been estimated on the basis of stream characteristics (natural, vegetated and meandering) at 0.045 $s/m^{1/3}$ (Ven te Chow, 1959) and the parameters B, m_1 and m_2 have been defined by a cross-section survey (Appendix B). This results in two unknown parameters that require calibration: the energy slope, s, and the location dependent h_0. It should be noted that the roughness and the slope are part of the same calibration factor, so a possible error in the estimate of n is compensated by the calibrated slope.
The prior ranges of the to-be-calibrated rating curve parameters are shown in Appendix E and were based on literature values and site inspection. The exponent b has a value of 1.59 in a rectangular channel, a value of 1.69 in a trapezoidal channel with side slopes of 1:1 and a value of 2.68 in a triangular channel (Luxemburg *et al.*, 2011). Therefore, the exponent b was given a prior range of 1 to 3. The coefficient a

was allowed to vary freely. The stage at zero flow, h_0, was first set at the lowest water level measured over the simulation period; this already gave rather good performance. However, this assumption is not entirely valid since it is possible there still is some flow at this lowest stage (possibly by seepage through bed forms), implying that the actual h_0 is lower, or the zero flow status is reached earlier (the lowest water level measured was present in a very dry period with a lot of evaporation), implying a higher actual value for h_0. To find the actual value for h_0 it was decided to apply a range of roughly 4 meter around the lowest stage observed and determine the value for h_0 by calibration. It was shown that the best model performance was obtained with a value for h_0 similar to the lowest value measured. Note that h_0 depends on the location; in Appendix E, h_{0-1} represents the h_0 applied to sub-catchment 1, whereas h_{0-2} corresponds with sub-catchment 2. The energy slope at station 2, equal to the bed slope under uniform flow, could be estimated using the DEM at 0.00017. However, since the DEM has a rather course spatial resolution (90x90 m), it was decided to determine the value by calibration, whereby the calibration range was defined between 0 and 0.002. The optimised slope appeared to be roughly two times higher than the estimated value (whereby it possibly compensates for the error in the roughness estimate).

3.2.3.2. Sensitivity rating curve to forcing

To check the sensitivity of the rating curve to the primary drivers E_{pot} and P, the influence of 50% more E_{pot} and a 50% less P on the rating curve has been tested. This influence is expressed as the 'potential evaporation elasticity of Q' and the 'precipitation elasticity of Q', which measure how the long term annual average of Q changes in response to a change in the long term annual averages of E_{pot} and P (Eq. 35 and Eq. 36). Here, the long term annual average is based on four years of data. An elasticity lower than 1 indicates inelasticity (discharge is less sensitive to the considered parameter), while an elasticity higher than 1 indicates elasticity (discharge is highly sensitive to the considered parameter) (Smith, 1776; Sawicz *et al.*, 2011).

$$E_{E_{pot}} = \frac{\partial Q}{\partial E_{pot}} \frac{E_{pot}}{Q} \quad (35)$$

$$E_P = \frac{\partial Q}{\partial P} \frac{P}{Q} \quad (36)$$

In which $E_{E_{pot}}$ (-) is the potential evaporation elasticity of Q, E_P (-) is the precipitation elasticity of Q, $\frac{\partial Q}{\partial E_{pot}}$ is the partial derivative of Q with respect to E_{pot} and $\frac{\partial Q}{\partial P}$ is the partial derivative of Q with respect to P, E_{pot} (mm/a) is the long term annual average potential evaporation, P (mm/a) is the long term annual average precipitation and Q (mm/a) is the long term annual average discharge.

3.2.3.3. Variability of the rating curve parameters

An important assumption in these methods is the invariability of the rating curve parameters over the chosen modelling period. This means that we assume a constant bed level and channel cross-section over a defined period. Obviously this assumption is not entirely valid. Sedimentation and erosion are continuous processes in alluvial rivers, hence there is not one single rating curve that describes the relation between H and Q over the years (e.g. Westerberg *et al.*, 2011). It is reasonable to assume that the largest variability is in the river bed elevation (due to varying bed forms) and not in the banks. Since the high flows are influenced rather by the total cross-sectional area, it is also reasonable to assume that the high flows are not as much affected by varying bed forms as the low flows. Since the goal of this research is to investigate the potential of modelling on stage rather than discussing rating curve stability it has been decided to accept the assumption of stable parameters over a time frame of five years. We see that the same assumption is made within the current practices of the Endau River catchment: rating curves are on average revised every five years. In order to implicitly account for changes in bed level and changes due to replacement of the gauge or reposition of a recorder, that affect the low flows, a modelling period was chosen in which no large variations in the low flow pattern of H could be distinguished. A preliminary recommendation is to investigate the influence of a variable value for h_0, whereby h_0 is recalibrated after each considerable flood, or when a shift in stage reading is observed.

3.2.4. Model forcing

The model input required consists of potential evaporation, precipitation and stage readings, all data is on daily basis. For the potential evaporation the corrected Penman open water evaporation has been used (Section 2.4.3.), calculated in geographical cells with 1.5 degree spatial resolution on the basis of ERA Interim, whereby for sub-catchment 1 and 2 different cells were used. For the precipitation in sub-catchment 1, gauge 4 has been applied and for sub-catchment 2 TRMM data. Finally, for both sub-catchments the observed water levels by JPS have been used.

The calibration period was selected in a way that no large variations in the low flow pattern of H could be distinguished, implying a rather constant h_0. The model was calibrated on four years of data, ranging from 2003 to 2007. The first year (2003-2004) was taken as spin-up time and not taken into account in the calculation of the objective functions in Section 3.2.5.1.

3.2.5. Calibration

Calibration is a process of parameter adjustment (automatic or manual), until observed and calculated output time-series show a sufficiently high degree of similarity (Beven, 2001b). For the Endau River catchment, a step wise calibration approach has been adopted, whereby first the rating curve parameters were determined, while all other parameters were kept free. This means that the fixation of the rating curve is in principle only based on the model structure and the model forcing. Model parameter values have no influence, since they are all free (within their prior range). During this process, it was found that h_0 could be well defined in the early steps. It was found easiest to first narrow the prior range of h_0, after which the optima of the other parameters became clearer as well.

Once the rating curve parameters had been fixed, the observed stage could be translated into a reanalysed hydrograph. Subsequently, from this reanalysed hydrograph, the range for the slow recession coefficient was obtained by plotting the recession on a semi-log scale (Appendix D). In the following step the conceptual model parameters were calibrated.

Each calibration step consisted of multiple (roughly 10) iterations, whereby 10.000 – 100.000 parameter sets were randomly generated using Monte Carlo sampling. After each iteration the parameter space was narrowed. In this subspace, again a large number of random parameter sets was generated.

3.2.5.1. Objectives

In order to assess model performance on different aspects of the system’s behaviour a multi-objective evaluation approach (Gupta *et al.*, 1998) has been adopted. The models were evaluated on the goodness of fit of the flow duration curve (Vogel and Fennessey, 1994; Smakhtin, 2001) (Eq. 37), the flow duration curve of the logarithm of the discharges (Eq. 38), the logarithm of the discharges (Eq. 39) and the discharges (Eq. 40), using the Nash-Sutcliffe Efficiency (Nash and Sufcliffe, 1970). The Nash-Sutcliffe

Efficiencies were all calculated for the entire modelling period, excluding the spin up period of one year. The models were calibrated as such to maximize the objectives, whereby a value of 1 implies a 100% fit, while a value lower than 0 implies the model performance is worse than when you would have taken the mean value of the observed flows (Savenije, 2009a).

Objective A:

$$(37) \quad NSE_{logFDC} = \frac{\sum\left(X_{logFDC,i} - Y_{logFDC,i}\right)^2}{\sum\left(Y_{logFDC,i} - \overline{Y_{logFDC}}\right)^2}$$

Objective B:

$$(38) \quad NSE_{FDC} = \frac{\sum\left(X_{FDC,i} - Y_{FDC,i}\right)^2}{\sum\left(Y_{FDC,i} - \overline{Y_{FDC}}\right)^2}$$

Objective C:

$$(39) \quad NSE_{logQ} = \frac{\sum\left(X_{logQ,i} - Y_{logQ,i}\right)^2}{\sum\left(Y_{logQ,i} - \overline{Y_{logQ}}\right)^2}$$

Objective D:

$$(40) \quad NSE_{Q} = \frac{\sum\left(X_{Q,i} - Y_{Q,i}\right)^2}{\sum\left(Y_{Q,i} - \overline{Y_{Q}}\right)^2}$$

Where X represents the modelled value, Y is the observed (here: reanalysed) value, the subscripts *logFDC*, *FDC*, *logQ*, *Q*, *i* and the overlie denote the flow duration curve of the logarithm of the discharge, the flow duration curve, the logarithm of the discharge, the discharge, the index and the average value.

Objective C and D are traditional objectives in the evaluation of hydrological models, which directly compare time-series with time-series. Objectives A and B are based on a hydrological signature rather than the exact time series. The flow duration curve, or the flow exceedance probability curve, indicates the catchment's ability to generate flows of different magnitudes (Yilmaz *et al.*, 2008). The Nash-Sutcliffe Efficiency of the flow duration curve (or the flow duration curve of the log of the flows) is therefore insensitive to the timing of the peaks or even missed rainfall events (as long as the statistics of the precipitation are well captured by the precipitation time series): hence this metric is a useful tool in scarcely gauged basins or basins with a relative uncertain series of forcing data, such as the Endau River catchment. By taking the FDC of the log of the flows, the emphasis is on the entire flow duration curve, while in case of the FDC of the flows, the main emphasis is on the high flows. Therefore, largest importance was given to objective A, directly followed by objective B. Objective C was considered third important and objective D fourth. In this order of importance, two exceptions were made. In finding the best value for h_0, main focus was given to objective C, since h_0 mainly influences the low flows; followed by A, B and D. In finding the best value for exponent b, main focus was given to objectives B and D, followed by A and C. This was done because, the exponent b has minor influence on the low flows, while its influence on the high flows is considerable, (Figure 10).

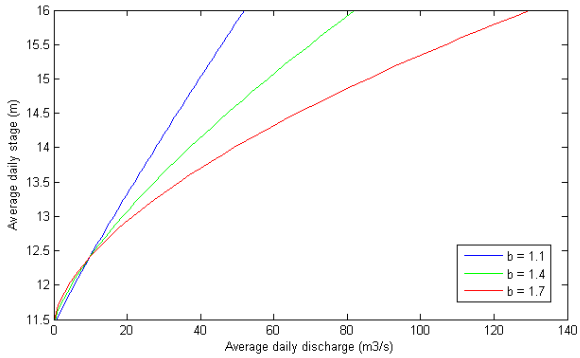


Figure 10 – The sensitivity of the rating curve to exponent *b*: *b* has a strong influence on the high flows, while it barely influences the low flows.

3.2.5.2. Narrowing the parameter range

The performance of each parameter value was visualised by various scatter plots, whereby the performance on the various objectives was plotted versus the parameter value. When the parameter was well defined, the parameter range could easily be narrowed on the basis of these scatter plots. Sometimes is was harder to distinguish the optimum range from the plots. Therefore an additional method was adopted in which objective thresholds were introduced, dividing the sample set in behavioural and non-behavioural models. These thresholds were defined as such to select the 10-25 best performing behavioural parameter sets and reject the others. These thresholds were model-, catchment-, step- and iteration- dependent. The parameters ranges were then narrowed according the parameter ranges of these behavioural models. Within this subspace, again 10.000 – 100.000 Monte Carlo samples were drawn. The number of samples used depended on the iteration: at an early stage the number of samples was lower, since the parameter space was still very wide and optimization was relatively easy. As the parameter space narrowed, more parameter set samples were required. The process of narrowing and re-sampling finished when no significant improvement could be found in the objective scores between the different iterations.

3.2.6. Validation

3.2.6.1. Model validation

To check whether the models performed right for the right reasons (Kirchner, 2006), the models have been validated on independent data, generated by running the model for a different period. The independent modelled hydrographs were assessed on the same four objectives A to D and additional objective E: the autocorrelation. The AC is a statistical property of the river flows, indicating the time scale characteristics of process components. Here, the lag-1 autocorrelation (Winsemius *et al*., 2009) was calculated, implying that the data series is compared with the same data series, shifted by one day. A high AC means that there is a small difference between two consecutive data points (Euser *et al*., 2013). This objective is represented by one value for the reanalysed hydrograph and one value for each modelled hydrograph. The evaluation criterion is calculated as Eq. 41, in which a higher value means a better performance.

Objective E:

$$(41) \quad F = 1 - \left|1 - \frac{AC(Q_{modelled})}{AC(Q_{reanalysed})}\right|$$

Where F (-) is the evaluation criterion, $AC(Q_{modelled})$ is the lag-1 autocorrelation of the modelled hydrograph and $AC(Q_{reanalysed})$ is the lag-1 autocorrelation of the reanalysed hydrograph.

To select the most balanced solution for each model, the Euclidean distance, γ, (e.g. Schoups *et al.*, 2005; Hrachowitz *et al.*, 2013) was introduced. γ is considered the ability of the model to reproduce the daily flow values in the validation period, denoted as Eq. 42. In this definition a lower Euclidean distance indicates a better overall performance.

$$(42) \quad \gamma = \sqrt{(1 - NSE_{logFDC})^2 + (1 - NSE_{FDC})^2 + (1 - NSE_{\log Q})^2 + (1 - NSE_Q)^2 + (1 - AC)^2}$$

Since not only the conceptual model parameters are validated on a different period, but also the rating curve parameters, it is important to choose a period in which these rating parameters are likely to be valid. It is assumed that a period in which the low flow pattern of *H* is similar, there is potential for the same rating curve parameters to be valid. For sub-catchment 2 no big jumps or trends in *H* can be distinguished (Figure 4b), therefore basically every period is assumed feasible. The period January 2009 to January 2012 has been chosen, whereby the first year functions as spin up time. In case of sub-catchment 1, a suitable period is less easy to find: after 2007 the behaviour of *H* considerably changed (Figure 3b). First it was tried to validate sub-catchment 1 for the same period 2009 to 2012. However, the rating curve parameters used over 2003 to 2007 gave very bad results in this period. In de validation period we observe unrealistically long periods with zero flow in the reanalysed hydrograph, implying that the value for h_0 used is likely too high. Additional periods have been tested, whereby January 2007 to January 2009 gave the least poor results.
It should be noted that this way of split-record validation does not proof anything, since it is very likely that the model will give satisfactory performance as the model is trained on data from the same location during calibration (e.g. Pokhrel and Gupta, 2011). Therefore we apply additional validation methods in Chapter 4 and 5.

3.2.6.2. Rating curve validation

The reanalysed rating curves for sub-catchment 2, found through calibration, have been validated by comparing them with additional discharge measurements, which were performed at the discharge station of sub-catchment 2 during a field visit (Appendix B). Note that the additional measurements are compared with a rating curve reconstructed for the period 2003-2007. The stage readings show a relatively constant pattern of low flows, indicating that no drastic erosion or deposition occurred over time. Moreover, during validation over the period 2009-2012 it appeared that the model with its corresponding rating curve parameters works quite well. Therefore this method is seen as a plausible approach.

3.3. Results and discussion

3.3.1. Rating curve parameter values

Figure 11 and Figure 12 show the scatter plots of the parameter value versus the model performances on objectives A to D. The plots of h_0 originate from a first iteration, so the entire prior parameter space has been taken into account. The plots for the other parameters originate from a second iteration, in which the

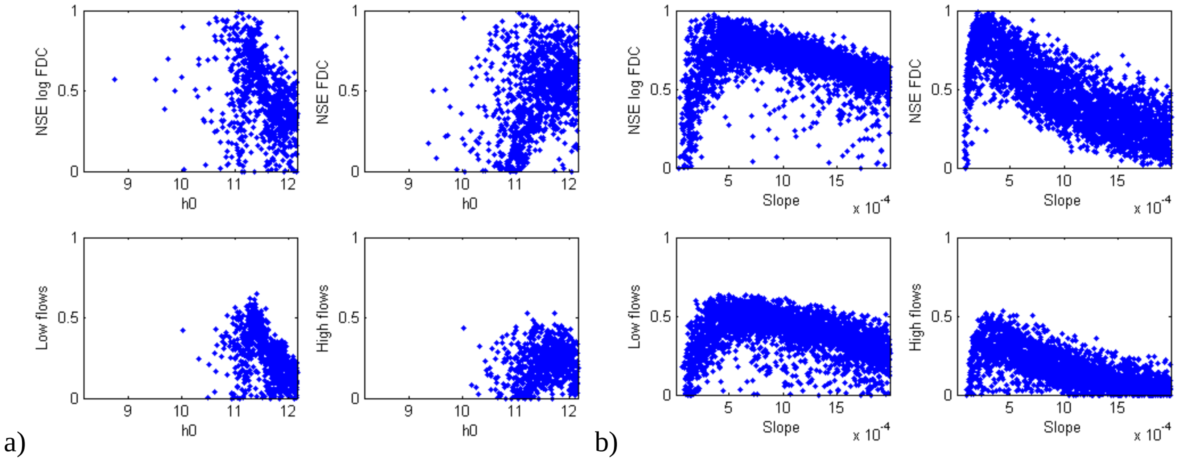


Figure 11 – Scatter plots in which all tested values for a) h_0 and b) the slope (of the 2-parameter definition), generated in the random sampling process are plotted versus their score on the four objectives described in Section 3.2.5.1. When the optimum parameter value is clearly visible, we call the parameter well defined.

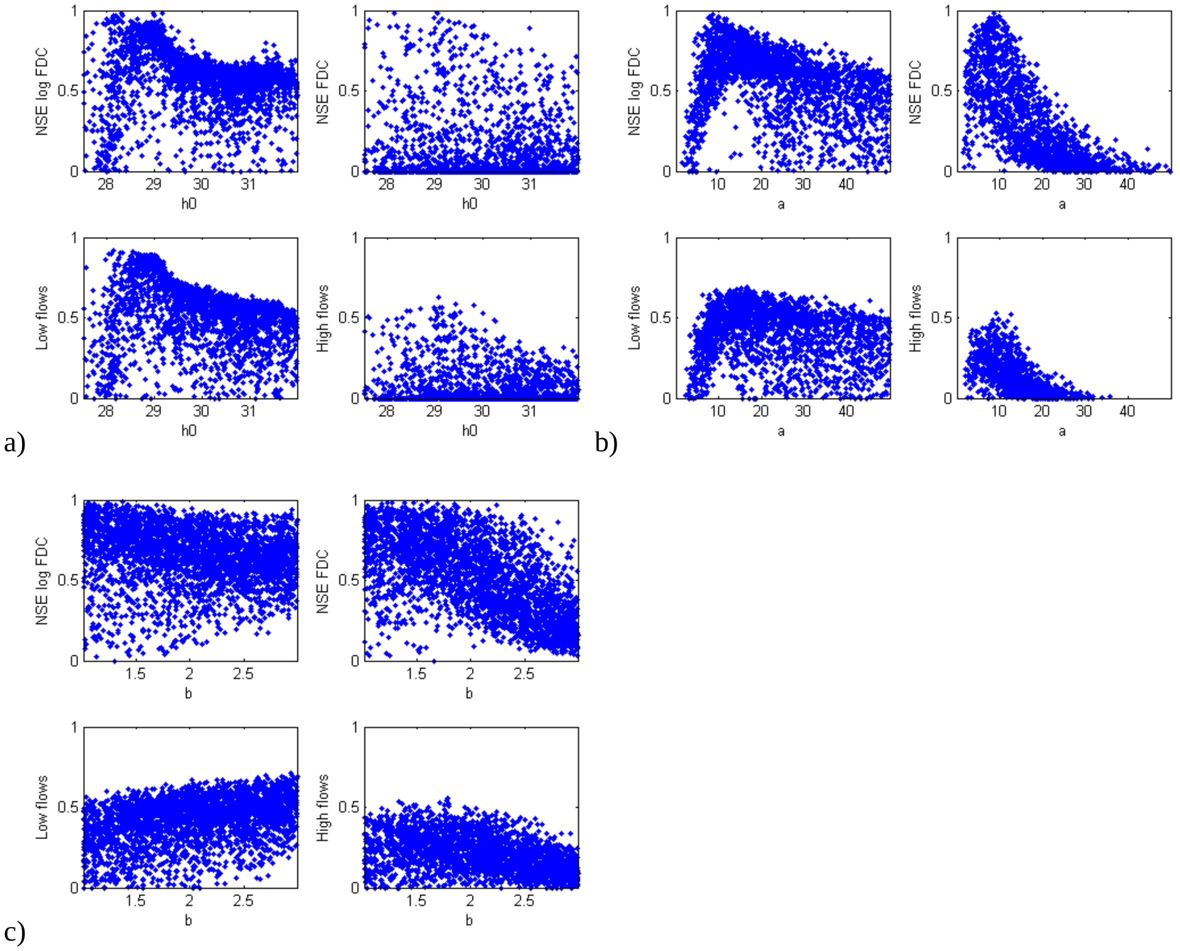


Figure 12 - Scatter plots in which all tested values for a) h_0, b) coefficient *a* and c) exponent *b* (of the 3-parameter definition), generated in the random sampling process, are plotted versus their score on the four objectives described in Section 3.2.5.1. When the optimum parameter value is clearly visible, we call the parameter well defined.

parameter space for h_0 has already been narrowed on basis of the first iteration. It can be seen that in the 2-parameter definition the optima are relatively well defined, despite of the fact that it is one of the first iterations. We also observe that all objectives optimise to the same optimum parameter range; there is no contradiction between objectives. In case of the 3-parameter definition the optimum for h_0 is again well defined, as becomes clear from Figure 12a. We also observe a clear optimum range for coefficient *a* (Figure 12b), while exponent *b* (Figure 12c) is less clearly defined. In case of exponent *b* we also observe conflicting objectives: the low flows perform best with a high value for *b*, while the other three objectives find their optimum in a lower range. As explained in Section 3.2.5.1., exponent *b* has minor influence on the low flows, while its influence on the high flows is considerable; therefore it was decided to narrow the parameter space after this first iteration according objectives A, B and D.

The posterior parameter distributions are given in Appendix F. The best model performances are obtained with values for h_0 close to the lowest water level measured. The calibrated slopes appear to be roughly two times higher than the estimated value through the DEM (whereby the slope possibly compensates for the error in the roughness estimate). The posterior range for *b* lies between 1.38 and 1.45, which is rather low when compared with literature values of 1.59 (rectangular channel) and 1.69 (trapezoidal channel) (Luxemburg *et al*., 2011). As stated before, the value for *b* is not so well defined. A way to improve the calibration of *b* could be to change the prior parameter range to 1.5 – 1.7 instead of 1 – 3, although a thorough literature study of possible values for *b* is recommended in advance.

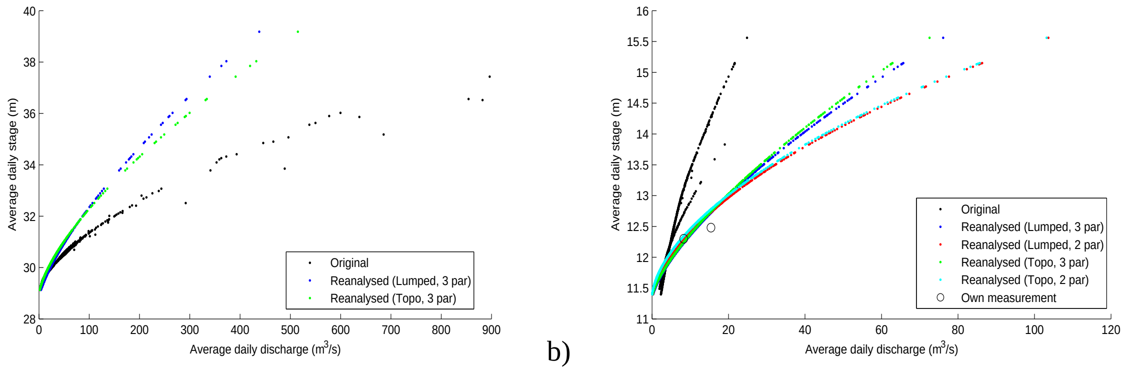
3.3.2. Reanalysed rating curves

The reanalysed rating curves found for sub-catchment 1 and 2, using different model structures (lumped and topography driven) and different rating curve definitions (only for sub-catchment 2) are shown in Figure 13 together with the original rating curves used in the same period (2003-2007).

We observe that the rating curves are relatively insensitive to model structure and relatively sensitive to the number of parameters used for the rating curve. We also observe that the rating curves, defined for the two model structures using the 2-parameter definition, show more similarity than the curves defined using the 3-parameter definition, see Figure 13b. We see for sub-catchment 1, that the beginning of the rating curve overlaps with the beginning of the original rating curve, which is considered to be the most reliable part, because it covers the range where most (if not all) discharge measurements have been done. At sub-catchment 2 we see reanalysed rating curves that are considerably different from the original curve, as expected from the data analysis in Section 2.4.1. When plotting the performed discharge measurements in the same figure, we observe that one measurement plots exactly on the reanalysed curve, while the second (less reliable) measurement plots slightly off the reanalysed curve.

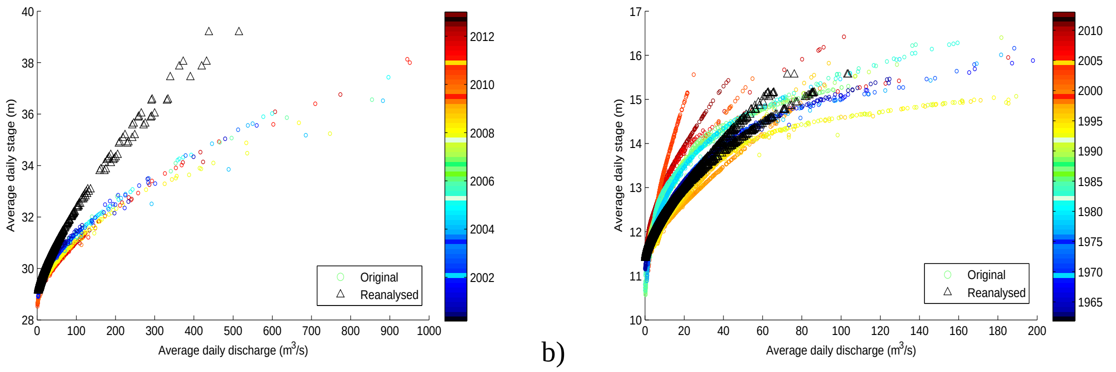
When comparing the reanalysed rating curves with the historical rating curves, the graphs of Figure 14 are obtained. We see that the reanalysed rating curves of sub-catchment 1 again overlap with the beginning of the historical curves and show an offset at the higher flows. The reanalysed curves for sub-catchment 2 fit within the range of historical hydrographs. The rating curves constructed using 2 parameters overlap with the very first rating curve constructed in 1962, which is considered a (if not the most) reliable rating curve. The rating curve at the beginning of a hydrological measuring campaign is often constructed by experienced hydrologists. The rating curves which are developed in a later stage may become less reliable, when they are constructed by so called 'desk hydrologists': hydrologists who have limited experience in the field.

Using these reanalysed rating curves together with the observed stage, reanalysed hydrographs were formed, of which 2 examples are shown in Figure 15. It can be seen that in case of sub-catchment 1, the



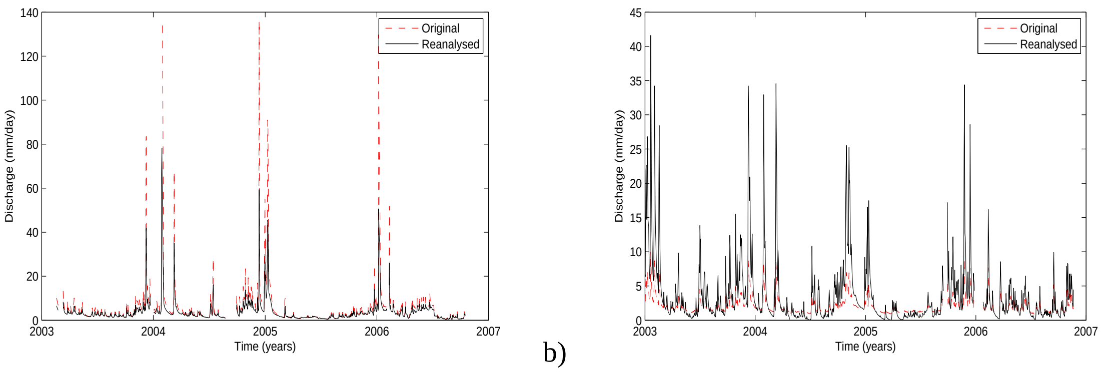
a) b)

Figure 13 - Original and reanalysed rating curves for the lumped and topography driven model structure for a) sub-catchment 1 and b) sub-catchment 2.



a) b)

Figure 14 - All historical original rating curves and all the reanalysed rating curves for a) sub-catchment 1 and b) sub-catchment 2.



a) b)

Figure 15 – Original and one of the reanalysed hydrograph for a) sub-catchment 1 and b) sub-catchment 2.

low flows are rather similar, while the high flows have almost halved in magnitude, in accordance with what was found in the rating curves. At sub-catchment 2 we see a relatively big change in hydrograph in both the low and the high flows.

3.3.3. Influence of forcing on rating curves

Figure 16a shows the influence of 50% more E_{pot} and 50% less P on the rating curve. For both situations, a clear shift to the left can be observed for the high flows, implying a lower discharge for the same stage reading. This makes sense: an energy constrained catchment (Section 3.3.7.) that receives less precipitation will generate less runoff and an energy constrained catchment that is exposed to a higher potential evaporation will have a higher actual evaporation and thus generate lower discharges.

It was found that an increase of 50% in E_{pot} leads to a decrease in Q of 27%. Hence, the elasticity, $E_{E_{pot}}$ is 0.54, implying an inelastic relation, thus a rather low sensitivity of Q to E_{pot}. A decrease in precipitation of 50% leads to a decrease in discharge of almost 80%, leading to an elasticity, E_P of 1.6, thus an elastic relation and therefore a high sensitivity of Q to P. From $E_{E_{pot}}$ and E_P it can be concluded that Q is not linear in P nor E_{pot}. An additional conclusion is that the primary driver P strongly influences Q. Since the rating curve is strongly determined by the water balance ($\frac{dS}{dt} = P - E - Q$), the rating curve is strongly driven by the primary driver P.

Figure 17 shows the actual evaporation and the storage in UR in sub-catchment 2 over time, according the lumped model forced by 100% E_{pot} and 150% E_{pot}, in which the maximum storage capacity of UR is similar in both models. To indicate wet and dry periods, the precipitation has been plotted as well. It can be seen that the higher potential evaporation causes the actual evaporation to be higher in periods with precipitation and lower in the periods with little or no precipitation. The former is caused by the fact that the Endau River catchment is energy constrained, implying that a higher potential evaporation leads to more actual evaporation in periods with enough moisture availability. The latter is a result of the higher evaporation flux, which causes the storage in the UR to drop faster and lower, visible in the middle graph of Figure 17. Under dry circumstances this can result in less actual evaporation despite the higher potential evaporation, due to a lower actual storage in UR (S_u), as indicated in Figure 16b.

3.3.4. Model parameter values

The posterior parameter ranges of the six developed models, as specified in Table 4, are given in Appendix F.

3.3.4.1. Lumped

Between the two lumped posterior parameter sets for sub-catchment 2 (based on the 2- and the 3-parameter definition), no large differences are observed. In sub-catchment 1 we find higher values for I_{max}, $S_{u,max}$ and P_{max} and a lower value for K_f compared to sub-catchment 2. All other parameters optimise to roughly the same range. These differences could be explained by the difference in topography. Sub-catchment 1 is hillslope dominated (Figure 2b and Table 1), with tropical rain forest as dominant vegetation, while sub-catchment 2 is wetland dominated, with grass land and crop land as dominant vegetation. Tropical rain forest has a higher interception capacity and deeper roots which comes with a larger maximum storage capacity in the root zone. The fast recession is faster (according to the calibration process) for the hillslope dominated sub-catchment than the wetland dominated sub-catchment, this could be due to a higher gradient, causing the potential energy and thus the kinetic energy to be higher as well. The higher percolation in the hillslope dominated sub-catchment could be caused by this same characteristic (higher gradient) and additionally by the fact that hillslopes may have soils with more macro pores, caused by for example a more intensive root zone, as a result of which percolation is easier and thus larger.

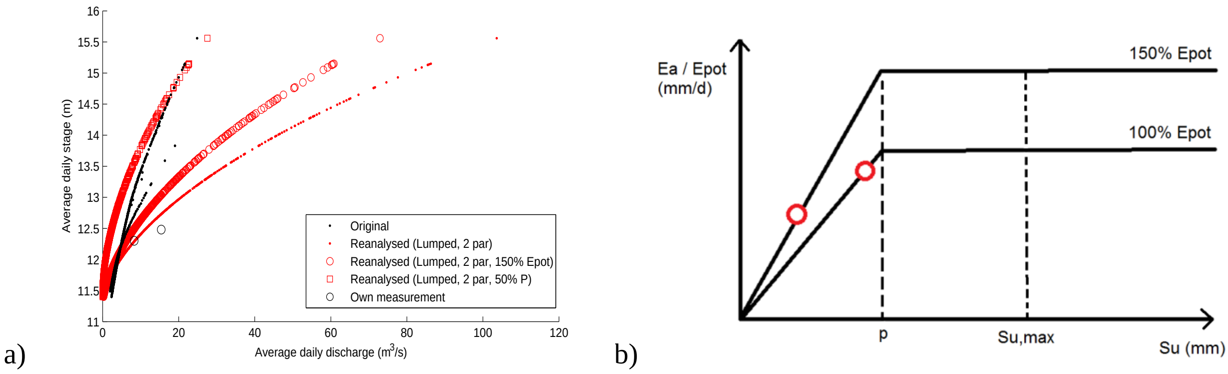


Figure 16 – a) The influence of a higher E_{pot} and a lower P on the rating curve and b) the conceptualisation of actual evaporation under influence of 100% and 150% E_{pot}. In a) it is observed that the influence of P is higher than the influence of E_{pot}. In b) we see that with the same storage conditions of S_u, the actual evaporation is always higher in the case with the higher E_{pot} . However, when the storage conditions of S_u are different, the actual evaporation from a model forced with a higher E_{pot}, could be lower, as indicated by the red circles.

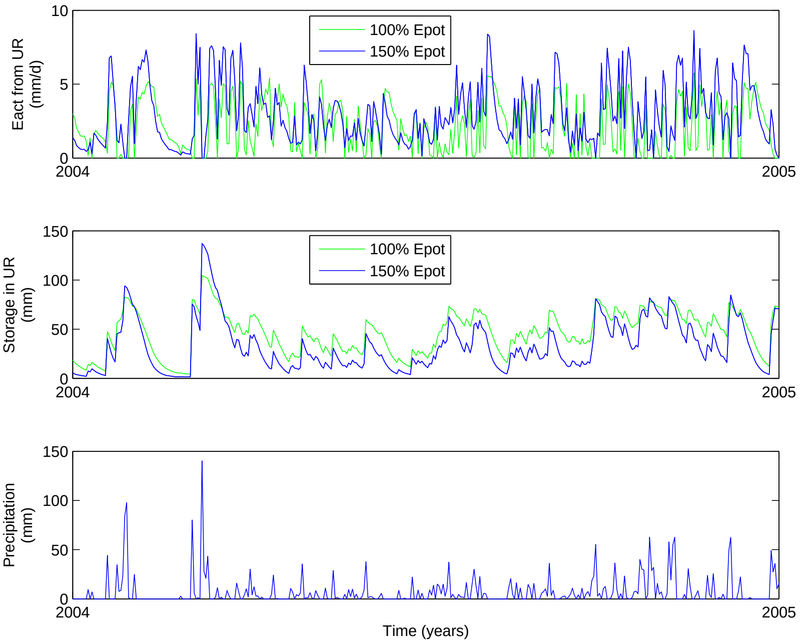


Figure 17 - Influence of a higher E_{pot} on E_{act} and the storage in UR together with the precipitation over the period 2004 – 2005.

3.3.4.2. Topography driven

In between the two topography driven posterior parameter sets for sub-catchment 2 (based on the 2- and the 3-parameter definition), the only parameter that gives considerably different results is $I_{max,W}$: a higher value is found for the 2-parameter definition. It could be that the (too) low value found using the 3-parameter definition is the result of over-parameterisation. In both sub-catchments we find that β_H is higher than 1, while β_W is lower than 1, implying that hillslopes only contribute when the soils are enough saturated, while there is always runoff from wetlands. This could be explained by the theory that pores only connect when the saturation level is high enough (e.g. Savenije, 2010). The topography based parameter ranges confirm the faster recession caused by hillslopes rather than wetlands: models where $K_{f,H} < K_{f,W}$, perform better than the reverse. It is found that $K_{f,H}$ and $K_{f,W}$ of sub-catchment 1 are shorter than those of sub-catchment 2. If the latter theory is valid, this faster recession may be caused by the steeper gradients present in sub-catchment 1. Last finding is that the $I_{max},_W$ and $S_{u,max},_W$ of sub-catchment 1 are considerably higher than found for sub-catchment 2. This may be caused by differences in wetland vegetation, or the fact that there is too little wetland in sub-catchment 1 to correctly define these parameters.

3.3.4.3. Lumped versus topography driven

When we compare the value for *D* of the lumped model with the value for D_H of the topographical model, we observe a large difference: D_H is roughly 0.3, while *D* is 0.7. Both parameters split the runoff in a fast component (*D or* D_H) and a component of preferential flow to the groundwater reservoir (*1-D or* $1\text{-}D_H$). This difference is in the first place possible, since the models are conceptualised differently, in which similar parameters can show different behaviour. Secondly, in the topography driven model, the recharge determined by *1-D* is the only flux that routs into the groundwater reservoir and this flux is only generated on the hillslopes. Meanwhile, in the lumped model, the flux indicated by *1-D* occurs over the entire catchment, additionally, the lumped model also has a percolation flux leading into the groundwater reservoir. Therefore it makes sense that the value for *1-D* is higher for the topographical model than for the lumped model.

When we compare the value of $S_{u,max}$ (the maximum storage capacity in the root zone) in the lumped model with the $S_{u,max,H}$ and $S_{u,max,W}$ in the topographical model, we observe in sub-catchment 1 that $S_{u,max}$ is within the range of the $S_{u,max,H}$, while in sub-catchment 2, $S_{u,max}$ is within the range of $S_{u,max,W}$. This can be explained by the fact that sub-catchment 1 is dominated by hillslopes, while sub-catchment 2 is dominated by wetlands.

3.3.5. Hydrographs and flow duration curves

Appendix H gives an overview of all six sets (as specified in Table 4) of modelled hydrographs and flow duration curves during calibration; Appendix G gives the corresponding objectives scores. We observe both qualitatively (visual inspection) and quantitatively (objective scores) that all models are well able to mimic the flow duration curve and the flow duration curve of the logarithm of the flows. In case of sub-catchment 1, objectives A, B and D give similar scores for the lumped and the topography driven model, while the lumped model has a significant higher performance on objective C (low flows). For sub-catchment 2, the two topography driven models and the two lumped models show similar performance on objectives A to D.

When we look at individual events we see that all models miss several peaks or generate peaks while no peaks have been recorded. Both observations can be due to the scarcity of the precipitation data set, the latter could additionally be caused by missing stage readings. We also observe over- and underestimation of peaks, possibly caused by the scarcity of the precipitation data set or a wrong model conceptualisation. Another reason for the underestimation of the peaks could be the influence of backwater effects. During high flow periods, the water transports all kinds of debris due to which some blockage may arise with backwater as a result. Backwater causes the water level at the location of the gauging station to be higher or lower while the same flow is present. When these stage readings are subsequently used to calculate discharge, too high discharges are obtained. The modelled underestimation of the peaks may thus in fact be an overestimation of the reanalysed peaks. During a field campaign, it has been observed that station 2 measures just upstream from a bridge, backwater influence at this station is therefore very likely to occur during high flows.

3.3.5.1. Decomposed hydrographs sub-catchment 1

In Appendix H also decomposed hydrographs are shown, whereby is distinguished between the different process components. When we look at the decomposed hydrographs of sub-catchment 1, we see in the lumped model that the fast outflow accounts for most of the peaks; only during extreme events the rapid surface overland flow (SOF) process becomes active. From the decomposed hydrograph it can be seen that SOF occurs at the beginning of 2004; apart from this moment, there are three other events with SOF. This corresponds with what was observed under very wet, but not extremely wet conditions in the field: no SOF was observed.

For the topography driven model we observe that the working of hillslope and wetland is rather similar: both processes are active during the relatively ‘dry’ season and the ‘wet’ season and have a similar recession slope. This behaviour could be explained by the fact that the Endau River catchment is a rather wet catchment, as a result, the moisture level in the root zone is always high enough for the hillslopes to generate runoff. Consequently, the peaks are also produced by the hillslopes and wetlands together, although the wetlands take the biggest share. The latter is striking, as sub-catchment 1 is hillslope dominated. The relatively low contribution of the hillslopes may be explained by the fact that hillslopes have a larger interception capacity and transpiration potential due to which more water is quickly evaporated. However, according to the topography driven model, the hillslopes evaporate (all evaporation fluxes included) on average only 0.2 mm/d more than the wetlands. Another reason is that a considerable part (+/- 70%, see Section 3.3.4.3.) of the precipitation that falls on hillslopes ends up as preferential flow which recharges the groundwater reservoir, as a result less water ends up as fast runoff and therefore the share of the hillslopes in the peaks is low.

The high flow peaks in the ‘wet’ season are equally well mimicked by the lumped and the topography driven model, however, when we look at the peaks during low flows, we see that the topography driven model gives better results; the most likely reason for this, is the addition of the wetland structure, which is able to generate peaks, regardless the moisture storage in the root zone, due to a β_W lower than 1 (Section 3.3.4.2.). Additionally the residual storage capacity in wetlands is only limited due to shallow soils and a shallow water table depth.

3.3.5.2. Decomposed hydrographs sub-catchment 2

In the lumped decomposed hydrographs of sub-catchment 2 we also observe that the fast outflow accounts for practically all peaks; SOF only occurs at the end of 2006. Direct reason for the less frequent SOF events is the higher threshold $S_{f,max}$. The model is probably optimised to this higher threshold, because the value for $S_{u,max}$ is much lower than in sub-catchment 1. In this way higher peaks can easily be modelled, since there is less storage capacity in the root zone: an extra low threshold causing SOF is apparently not necessary.

In the topography driven models for sub-catchment 2, we observe that the wetland share is even larger than in sub-catchment 1, as a result of the wetland domination. The same explanations are valid as described before. Similar to what was found in sub-catchment 1, we see the peaks in the ‘dry’ season are relatively better represented by the topography driven models, the most likely reason for this is again the addition of the wetland structure.

3.3.6. Model validation

Appendix I presents an overview of the hydrographs and flow duration curves during validation, Appendix G gives the corresponding scores on the individual objectives and the Euclidean distance, γ. The models for sub-catchment 2 look rather good, while the models of sub-catchment 1 seem not very well able to mimic the reanalysed hydrographs. The latter is likely caused by the fact that the parameter h_0 could not be transferred to this new period: we observe a clear shift between the reanalysed and modelled flows, which could be caused by a too high value for h_0. Another possible cause is that the stage data is flawed, since after 2007 the behaviour is considerably different than before and a lot of data is missing.

Validation of sub-catchment 2 gives reasonable scores on all objectives. The topography driven models score better than the lumped models on objectives B and D, objectives that focus on high flows. This could be explained by the working of the wetland. In this landscape class, the dominant runoff generation mechanism is saturation excess overland flow, a lateral process that shows a fast response to precipitation and thus account for the sharp and high peaks. Also better performance is obtained on objective E, which is a statistical property of the river flows, indicating the time scale characteristics of process components. This could be caused by the fact that the topography driven model is able to generate different processes (e.g. shallow sub surface drainage, saturation excess overland flow) with different time scales.

γ gives very bad results for sub-catchment 1, due to the low scores on the individual objectives. In sub-catchment 2, γ gives best results for the topography driven models. The most balanced solution (i.e. the best score for γ) of the two topography models created, is the one whereby two parameters were used to define the rating curve (Appendix G). This is most likely caused by the fact that the parameter in the 2-parameter-defnition were better defined that those in the 3-parameter definition and therefore give a better representation of reality.

3.3.7. Catchments on Budyko

Out of curiosity and to perform a sanity check, the sub-catchment positions on the Budyko curve (Budyko, 1974) have been determined and plotted, shown in Figure 16b (blue and red markers). We see that both sub-catchments end up at roughly the same location on the segment representing energy constrained conditions. Wet and humid catchments, such as the Endau River catchment, are often energy constrained rather than moisture constrained, thus this outcome is not unrealistic. It is striking however, that both sub-catchment 1 and sub-catchment 2 plot on roughly the same location on Budyko, while the sub-catchments show obvious dissimilarities in e.g. landscape. We would for example expect sub-catchment 1 to have a larger runoff coefficient due to the dominance of hillslopes, however, this is not reflected here. The similar positions could be the result of the stage calibration, which is strongly driven by the water balance: it allows the discharge to be adjusted to the P and E_{pot}. Therefore the location on Budyko is no proof of our rating curve to be correct.

When sub-catchment 2 is forced by a higher E_{pot} or a lower P, the location on the Budyko curve (green markers) moves away from the energy constrained segment. This makes sense, as the dryness index as well as the evaporative index increases as a result of a higher E_{pot}/lower P and a lower Q.

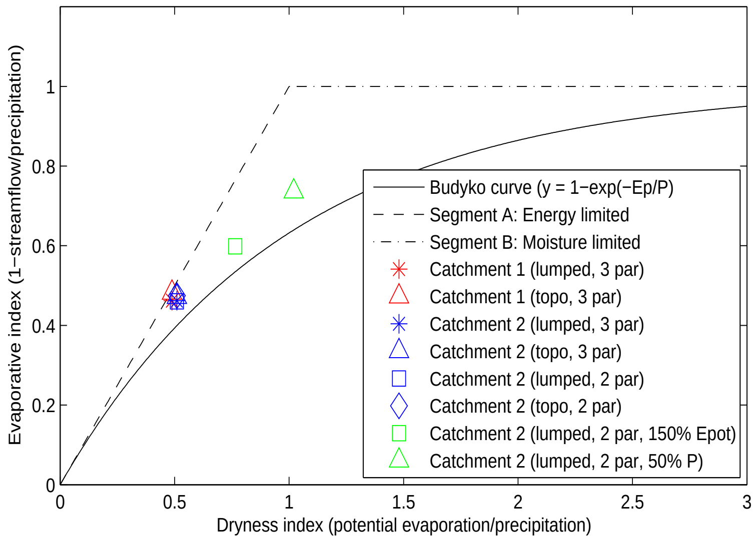


Figure 18 - The location of the sub-catchments on the Budyko curve according the reanalysed data.

3.4. Conclusion

On the basis of the results in Chapter 3, we can confirm hypothesis 1: Calibrating a rainfall-runoff model on stage does reduce discharge uncertainty in poorly gauged basins. Calibration on stage gives plausible results in the poorly gauged Endau River catchment and has potential for being applied more widely.
The most important findings with respect to calibration on stage are that the rating curve parameters are well defined in case of the 2-parameter definition and optimise to values that correspond to the physical property they represent. This implies that the addition of rating curve parameters does not increase the risk of over-parameterisation. In a sensitivity analysis, it was found that the reanalysed rating curves are insensitive to model structure and not very sensitive to the potential evaporation while a strong sensitivity was found to the precipitation forcing. The models that were found through calibration on stage are well able to mimic the reanalysed hydrographs, while it was impossible to find a good fit for the original 'observed' hydrographs. In sub-catchment 2, we observe that the reanalysed rating curves overlap with the discharge measurements done during the field visit. Additionally, an overlap is observed between the reanalysed rating curves of sub-catchment 2 and the first historical rating curve, which is considered the most reliable rating curve. In sub-catchment 1, we observe a clear overlap with the initial section of the original rating curve, which is considered the most reliable part.
After the rating curves had been established, the model hypotheses on catchment behaviour were tested by conventional methods. The most important findings from this analysis are that during calibration, the lumped and topography driven models show the same performance, while during validation on a different period, the topography driven models give the most balanced solutions. It was found that models, calibrated on reanalysed hydrographs based on the two-parameter definition, give the best results in a different period.
To improve the general applicability of this approach in other catchments, it is recommended to use the 2-parameter rating curve definition, in which both parameters are well defined and with which the most stable models are found. This method requires cross-sectional information, therefore it is recommended to always perform an adequate survey of the gauging site (profile, h_0, slope). In this research the cross-sectional area has been assumed to be trapezium-shaped. This is no natural cross-sectional shape,

therefore it is recommended to test the performance of a power function and a cosine function to describe the natural shape of the cross-sectional area. Another assumption in this study is the constant value for h_0, however, erosion and sedimentation are continuous processes. It is therefore advised to keep h_0 dynamic: recalibrate h_0 after each considerable flood, or when a shift in stage readings is observed. Since the rating curve is strongly driven by the water balance and the water balance in its turn by the precipitation, it is recommend to perform a solid rainfall analysis.

To check whether the models perform right for the right reasons, it is recommended to validate the models on independent data, preferably by application to a different catchment in a similar climate (as shown in Chapter 4). To validate the reanalysed rating curve it is recommended to carry out at least two discharge measurements during a short field visit.

CHAPTER 4

Hypothesis two: Topography driven rainfall-runoff models are better transferable than lumped rainfall-runoff models

4.1. Introduction

The split-record validation, as described in Section 3.2.6.1. does not proof anything, since it is very likely that the model will give satisfactory performance in a different period, as the model is calibrated on data from the same location (e.g. Pokhrel and Gupta, 2011). In this chapter the models are therefore validated by transfer to a close by sub-catchment. In this way, it is tested whether the models are able to reproduce the hydrologic response in a catchment they have not been explicitly calibrated on (Gao *et al.*, 2013) This type of validation yields insight in the process realism and the predictive power of a model. When the validation gives good results one could see this as a legalisation for application of the model to a similar ungauged catchment.

Additional implication of good validation results, is the implicit validation of the calibration on stage approach: the model is tested on its capability to reproduce the hydrologic response of an independent catchment, whereby the hydrologic response was derived from a rating curve that was calibrated on stage. Thus, when an independent model is able to reproduce this response, this implicitly validates the approach of calibration on stage.

4.2. Methods

The behavioural calibrated parameter sets found for the lumped model and the topography driven model for sub-catchment 1 have been transferred to sub-catchment 2 and vice versa. This resulted in 6 newly modelled hydrographs. These newly modelled hydrographs have been compared with the existing reanalysed hydrographs of the two sub-catchments. When transferring to sub-catchment 2, the hydrographs were thus compared with both reanalysed hydrographs of sub-catchment 2 (based on the 2- and the 3-parameter definition). An overview of the transfer process is given in Table 5. In the transfer process, only the model parameter sets were transferred, all other (location dependent) factors, were kept the same, being: the rating curve parameters, the forcing and in case of the topography driven model the parameters describing the part of the catchment classified as wetland (W) and hillslope (H) (Section 2.3., Table 1).

The model outcome was evaluated on the four individual objective functions A to D (Section 3.2.5.1.) and the Euclidean distance, γ, calculated according Eq. 43. Note that here, the Euclidean distance does not include the term on objective E, the autocorrelation.

$$\gamma = \sqrt{(1 - NSE_{logFDC})^2 + (1 - NSE_{FDC})^2 + (1 - NSE_{\log Q})^2 + (1 - NSE_Q)^2} \tag{43}$$

In which NSE_{logFDC} (-) is the NSE on the flow duration curve of the logarithm of the discharges (Eq. 37), NSE_{FDC} (-) is the NSE of the logarithm of the discharges (Eq. 38), $NSE_{\log Q}$ (-) is the NSE of the logarithm of the discharges (Eq. 39) and NSE_Q (-) is the NSE of the discharges (Eq. 40) (Section 3.2.5.1.).

4.3. Results and discussion

Appendix G shows the individual objective scores and the value for the Euclidean distance, γ, and Appendix J shows the hydrographs and flow duration curves during validation on a different sub-catchment. The differences in percentage between the calibration and different sub-catchment validation are shown in Table 5. The most stable solutions during transfer are obtained with the topography driven models trained on sub-catchment 2.

4.3.1. Sub-catchment 2 to sub-catchment 1

Appendix J shows that the transfer of the lumped models from sub-catchment 2 to sub-catchment 1, results in an overestimation of the peaks and an underestimation of the low flows. In Table 5 this over- and underestimation results in a large decrease in performance. The hydrographs and flow duration curves of the transferred topographical models look relatively good, which is confirmed by the relatively low decreases in performance in Table 5. The topography driven models show good performance on especially the flow duration curve and the logarithm of the flow duration curve simulation (objective A and B). This illustrates the robustness of the topography driven models in the flow frequency simulation. Also the integrated Euclidean distance shows a better transferability of topography driven models than lumped models.

The reason for this clear difference in performance is that topographical models distinguish between different topographical classes, while a lumped model does not. A lumped model has parameter values which represent more or less the average of the spatially distributed parameters values within this sub-catchment. When such a lumped model is transferred to a catchment where the spatial distribution, is different (e.g. caused by differences in topography between the two considered catchments), the model parameter values are no longer valid. This problem is largely overcome when different topographical classes are distinguished, as done within the topography driven model. It is assumed that these different topography classes have a similar working in different catchments (since the classifications reflect different runoff mechanisms and land cover) and that the parameters of a topographical model are specific for the topographical class they describe in a certain climate; hence the parameter values of the topographical model have much more physical meaning. When a topography model is transferred to a new catchment, the distribution of the topography classes is adjusted to the distribution in the new catchment, so that the model is a more reliable representation of what actually happens in the catchment. In a lumped model this cannot be done.

4.3.2. Sub-catchment 1 to sub-catchment 2

In Appendix J we see that transferring the lumped model of sub-catchment 1 to sub-catchment 2 leads to an overestimation of the low flows for both reanalysed hydrographs (3-parameter and 2-parameter definition). The high flows of the modelled hydrograph correspond rather well with the high flows of the reanalysed hydrograph based on the 3-parameter definition. The high flows of the reanalysed hydrograph based on the 2-parameter definition are considerably higher and underestimated by the model.

Appendix J shows that the transfer of the topography driven model of sub-catchment 1 to sub-catchment 2 gives better results than the lumped models. The model is well able to mimic the reanalysed hydrograph of the 2-parameter definition, while with the 3-parameter definition the high flows are mostly overestimated and the low flows slightly underestimated. This is likely caused by a wrong estimate of the exponent b (not well defined in calibration), due to which the reanalysed high flows are lower. Table 5 confirms that the most stable solution during transfer is again the topography driven model (with respect to the 2-parameter definition reanalysed rating curve). We see again good performance on especially the flow duration curve and the logarithm of the flow duration curve simulation (objective A and B). This illustrates the robustness of the topography driven models in the flow frequency simulation. The integrated Euclidean distance also confirms the best transferable model is the topography driven model: this model is well able to mimic the hydrologic response of the 2-parameter definition.

Table 5 – Differences in percentage between objective scores during calibration and objective scores during validation on a different sub-catchment. Green indicates an increase, or a decrease in performance of less than 10%, yellow indicates a decrease between 10-25%, orange indicates a decrease between 25%-40% and red implies a decrease higher than 40%. NB C2, 3 par denotes: sub-catchment 2, 3-parameter definition.

Model type	From	Transferred to	Δ objective scores				
			A	B	C	D	γ
Lumped	C1, 3 par	C2, 3 par	-15%	-6%	-10%	-37%	39%
	C1, 3 par	C2, 2 par	-22%	-10%	-14%	-34%	42%
Topo	C1, 3 par	C2, 3 par	-3%	-2%	-6%	-85%	59%
	C1, 3 par	C2, 2 par	-1%	5%	28%	-54%	23%
Lumped	C2, 3 par	C1, 3 par	-16%	-11%	-81%	-13%	64%
	C2, 2 par	C1, 3 par	-37%	-23%	-139%	-53%	125%
Topo	C2, 3 par	C1, 3 par	-3%	-18%	-25%	5%	20%
	C2, 2 par	C1, 3 par	-2%	-12%	-36%	14%	21%

VERY GOOD	GOOD	BAD	VERY BAD

4.4. Conclusion

The most important finding of this chapter is that topography driven rainfall-runoff models are better transferable than lumped rainfall-runoff models. These models show good performance on especially the flow duration curve and the logarithm of the flow duration curve simulation (objective A and B), which illustrates the robustness of the topography driven models in the flow frequency simulation. Also the integrated Euclidean distance shows better transferability of topography driven models as compared to lumped models. Therefore, hypothesis 2 is confirmed: Topography driven rainfall-runoff models are better transferable than lumped rainfall-runoff models. This implies that the process realism and predictive power of the topographical models is higher than of the lumped models.

A second finding is the validation of the stage calibration approach: the topography driven models were capable to reproduce the hydrologic response of an independent catchment, whereby the hydrologic response of the independent catchment was derived from a rating curve that was calibrated on stage. This adds to the credibility of the proposed method of calibration on stage. It was found that the transferred models were better able to reproduce the hydrologic response based on the 2-parameter definition than on the 3-parameter definition. Therefore it is (in line with the conclusion in Section 3.4.) recommended to use the 2-parameter definition which requires a solid cross-section survey.

CHAPTER 5

Hypothesis three: Linking a rainfall-runoff model to a steady state salt intrusion model offers mutual model improvement

5.1. Introduction

5.1.1. Motivation

Many people in the world live in coastal areas near estuaries, as estuaries function as a source of food and as a transport link between river and sea. Additionally, the lands adjacent to estuaries commonly have the perfect characteristics for agriculture: soils are fertile, the land is flat, and fresh water is available (Savenije, 2012). Salt intrusion can be a problem for agriculture and drinking water in these regions. Models are required to understand the processes that influence the salt intrusion and the possible effects of environmental changes or human impacts. A lot of research was done on one-dimensional salt intrusion models (e.g. Van der Burgh, 1972; Fischer, 1974; Sanmuganathan, 1975; Prandle, 1985; Savenije, 1986,1989,2003; Van Rijn, 2011). An important parameter within these models is the river discharge, Q_f (e.g. Nguyen *et al.*, 2006). In the most favourable case that the river discharge is actually measured upstream of the estuary mouth, it is done at a point where tidal influence is no longer present. This point is often only reached at dozens of kilometres upstream of the estuary mouth. Contributing streams downstream of this point are not taken into account, resulting in an underestimation of the actual discharge. Therefore, the uncertainty error in river discharge can be rather high, which directly translates in an uncertainty in the modelled salt intrusion length.

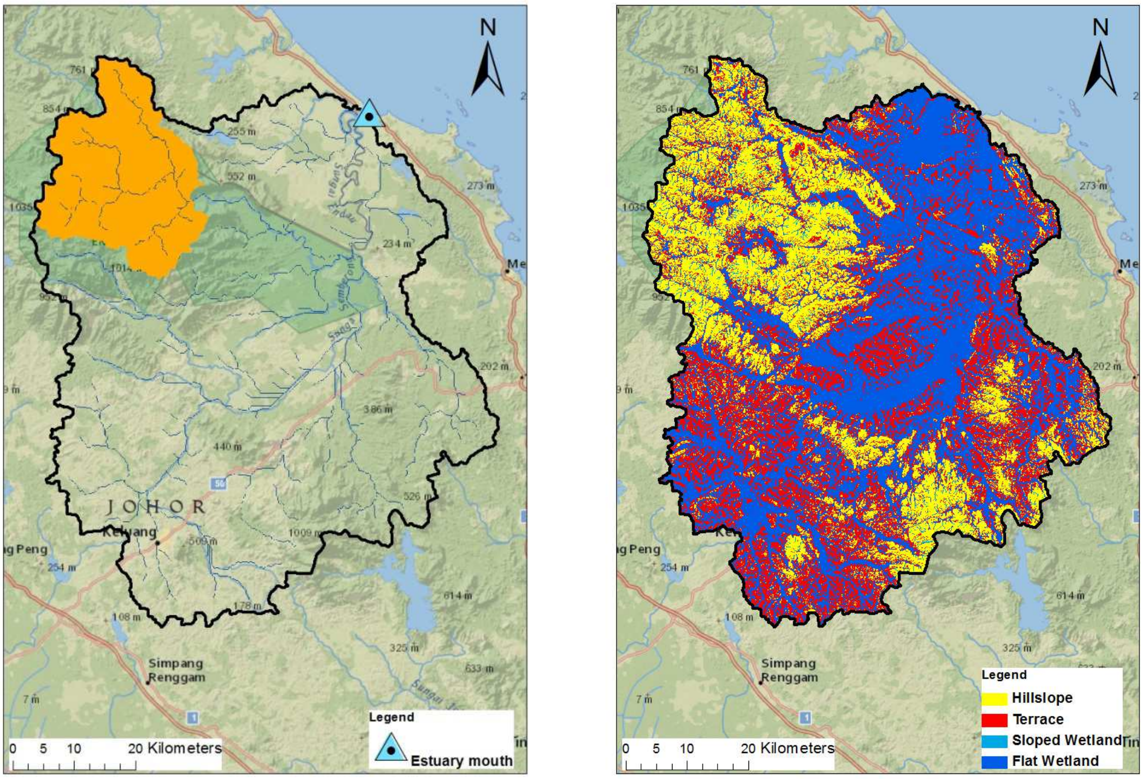
Another main research topic of hydrologists is the understanding of catchment hydrology to predict e.g. river discharges, by application of for example a conceptual rainfall-runoff model, as done in this research.

Now as it can be noted, there is a clear connection between a predictive salt intrusion model and a rainfall-runoff model: both topics deal with a river discharge, either as the model input or outcome. This chapter aims at highlighting the vast variety of mutual opportunities which the coupling of these two fields of study offers. To illustrate, a worked out example is given of the estuary of the Endau River catchment in South-East Malaysia.

5.1.2. Worked out example

On the 28th of March in 2013, the salt intrusion length was measured for the Endau River estuary, as reported by Nijzink (2013). To calculate this same intrusion length using a steady state salt intrusion model, the fresh water discharge of the river is required as input parameter for the model. However, on the day the salt intrusion length was measured, only one discharge station was active, corresponding with the discharge coming from the orange shaded area in Figure 19. One can imagine that the intrusion length calculated using only this discharge will be highly overestimated, since the fresh water discharge is highly underestimated. A rainfall-runoff model that simulates the river discharge at the mouth of the estuary could be the solution to this problem.

The other way around, the required discharge (Q_f) to generate a certain salt intrusion length, can be back calculated using the steady state salt intrusion model. Subsequently, the value for Q_f can be compared with the discharge predicted by the rainfall runoff model and in this way facilitate the validation of the upscaled model.



a) b)

Figure 19 – a) The upstream area flowing into the estuary with the gauged sub-catchment shaded in orange and b) the full catchment classification into hillslope, terrace, sloped wetland and flat wetland.

5.2. Methods

5.2.1. Discharge from salt intrusion model

The discharge from the salt intrusion model was calculated with a steady state model based on the theory developed by Savenije (2012). Savenije (2012) states that the shape of alluvial estuaries can be described by exponential functions. The longitudinal distribution of salinity, tidal damping, phase lag, wave celerity and velocity can be computed on the basis of topography, tidal condition, fresh water discharge, friction, the salinity and the tidal amplitude at the estuary mouth. The most important equations are shown in Eq. 55 to Eq. 65 in Appendix L.

The unknown parameters in these equations have been determined through various methods: the parameters υ, C, E and P_t were based on a tidal dynamics model (Cai *et al.*, 2012; Savenije, 2012) and the parameters S, S_0, S_f, a, b, h, x, A_0, B_0, h_0 and T were directly obtained from field measurements. With the knowledge of these parameter values, the salt intrusion length, L was calculated. To check the sensitivity of Q_f to L, a sensitivity analysis was performed.

5.2.2. Discharge from rainfall-runoff model

5.2.2.1. Model selection

The model that has been developed for one of the sub-catchments and performed best during calibration, validation (on a different period and different catchment) was selected for upscaling to the full Endau River catchment to simulate the discharge at the estuary mouth. It was assumed that when a model performs well on two sub-catchments within the full catchment, (while it was only explicitly trained on one of these), the upscaled model should be able to simulate the hydrologic response of the full Endau River catchment as well. Since a model is used that was calibrated on stage, good validation results during coupling implicitly confirm the potential of calibration on stage as well. Appendix G shows that the best model is the topography driven model for sub-catchment 2, calibrated on the reanalysed hydrograph based on the 2-parameter definition for the rating curve.

As explained in Section 3.2.5., not one best parameter set has been selected, but a set of behavioural parameters sets. The upscaled model was run for all of these behavioural parameter sets over the period 2003 – 2013. This resulted in a set of time series of daily flow values at the estuary mouth. To have one value to compare with the discharge from the salt intrusion model, the typical discharge was calculated as the average of all behavioural models.

It should be noted that the parameters describing the part of the catchment classified as wetland (*W*) and hillslope (*H*), were set according the classification of the full Endau River catchment. The full Endau River catchment is classified as 75% wetland and 25% hillslope, as visualised in Figure 19 and quantified in Table 1.

5.2.2.2. Forcing

For the precipitation forcing a representative time series has been used originating from TRMM data. The TRMM cell used was selected as follows: the long term annual average of all TRMM cells covering the Endau River catchment was calculated. Subsequently, the average of all these long term annual averages was taken. The TRMM cell that has a similar annual average and is simultaneously located most centrally was selected. The long term annual average was calculated in two ways: in the first method only the four years of actually used data were included; in the second method all ten years of available data were used. Both methods resulted in the same decision for the representative TRMM cell. For the potential evaporation, again the corrected Penman open water evaporation was used (Section 2.4.3.), originating from the geographical cell covering the largest part of the catchment.

5.2.2.3. Travel time

As specified in Section 3.2.1., no routing function was included in the model structure, since timing was before considered unimportant. However, now we are comparing absolute discharges, at a point where channel routing without doubt plays an important role (Savenije, 2009a), since the travel distance to the estuary mouth is considerably longer than within the individual sub-catchments. Therefore, it is very likely that the actual discharge at the mouth is delayed by a few days, implying that the discharge calculated by the model is in fact the discharge that will arrive a few days later. To get a rough estimate of the delay, the time was calculated it takes for a water drop to travel from the furthest part in the catchment to the estuary mouth (as the crow flies). The maximum distance a drop has to travel is roughly 80 km. The average flow velocity, based on observed flow velocities within the catchment, was estimated at 0.5 m/s. The travel time was therefore calculated as roughly 2 days. The total discharge was delayed by this time.

We realize this is only a very rough estimate, whereby we do not distinguish between different delays for different fluxes, however we do think the delayed discharge is more representative than the non-delayed discharge. It is highly recommended to apply more advanced methods that account for the delay to improve the general applicability in other catchments.

5.3. Results and Discussion

5.3.1. Modelled discharges

The salt intrusion length, L, measured on the 28th of March is roughly 29 km. The steady state salt model gives a corresponding average daily river discharge of roughly 80 m^3/s. This discharge is highly sensitive to the measured salt intrusion length: using Eq. 54 – 64 (Appendix L) it was calculated that a 10% decrease/increase in L results in a 25% increase/decrease in Q_f, implying an elasticity of 2.5.
The average daily discharges predicted by the upscaled rainfall runoff model, on and five days prior to the day of measuring are shown in Table 6. Figure 20 shows the fresh water discharge over time at the estuary mouth from the rainfall-runoff model (black and green graphs) with respect to the predicted fresh water discharge by the salt model (red dot). The black graphs represent the non-delayed discharge, whereas the green graphs represent the 2-day delayed discharge. Although (or thanks to?) the considerable amount of uncertainty present, the results show that the two modelled discharges are within the same order of magnitude. If we assume that the estimated delay is correct, the discharge according to the rainfall-runoff model is roughly 102 m^3/s, which gets already closer to the 80 m^3/s predicted by the steady state salt intrusion model. Moreover, the delayed hydrograph does not take account of the wave attenuation, which would reduce the modelled value of 102 m^3/s even further. This adds to the credibility of the applied methods: the upscaling seems reliable. Since the upscaled model originates from a model that was calibrated on stage, the calibration on stage method is implicitly validated as well. However, we only compare one moment in time, which is not very convincing. This method would gain more confidence when longer time series were compared. In this way it is also possible to not only verify the absolute value, but also the trend.

5.3.2. Uncertainty

In addition to the fact that comparing one moment in time is not very convincing, there is also a lot of uncertainty present in the models that give a prediction for that very moment. Both the salt model's discharge and the rainfall-runoff model's discharge are sensitive to various error-sources. Important issues that may have a large effect on the rainfall-runoff model's discharge are upscaling of the model and the effect of ocean backwater in the estuary. Due to the widening of the estuary the fresh water flood wave is attenuated even more, due to which peaks are flattened out (Savenije, 2012). In the upscaling process we deal with heterogeneity and variability in the catchment, where the term heterogeneity is commonly used for media properties (such as porosity) that vary in space, whereas variability is typically used for fluxes (such as evaporation) or state variables (such as soil moisture) that vary in space and/or time (Blöschl *et al.*, 1995). The heterogeneity and variability are partly accounted for by distinguishing between the topographical classes, but this may not be sufficient. Disparity in e.g. geology or land use may be present in between the wetlands of the sub-catchments and the wetlands of the lands closer to the delta, resulting in different behaviour and thus a different hydrologic response. Another factor imposing a large uncertainty is the lumped precipitation forcing: in reality the precipitation has a highly heterogenic character. To better account for heterogeneity in precipitation one could apply distributed precipitation as

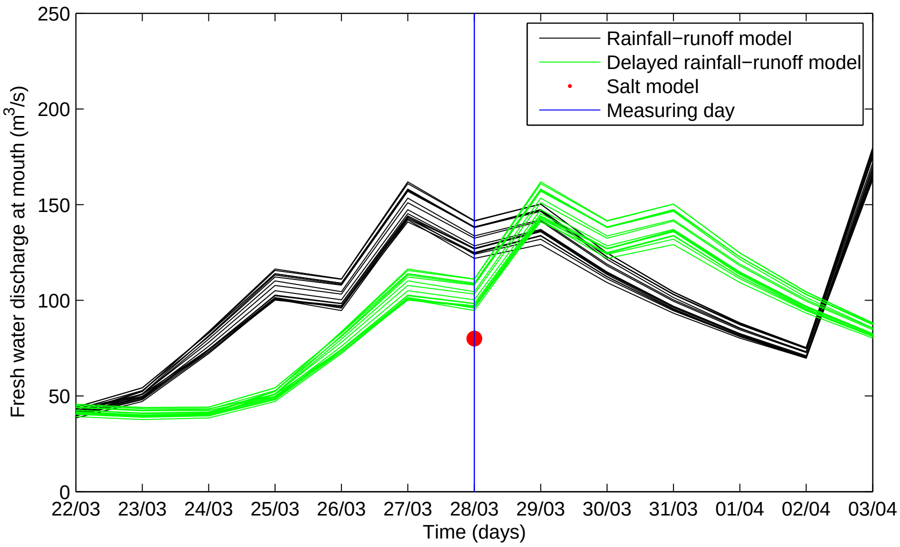


Figure 20 - Discharge at the estuary mouth according the steady state salt intrusion and upscaled rainfall-runoff model. The black graphs represent the non-delayed discharge, while the green graphs include a delay of 2 days.

Table 6 - Average daily non-delayed discharge at estuary mouth according the upscaled rainfall runoff model.

Date	Average daily discharge (m^3/s)
23-3-2013	51
24-3-2013	77
25-3-2013	107
26-3-2013	102
27-3-2013	150
28-3-2013	131

input, while keeping the model parameters lumped (e.g. Fenicia, 2008).

Concerning the uncertainty within the discharge predicted by the salt intrusion model, we can address several uncertainty factors as well. In the way the salt intrusion model was used here, the most relevant uncertainty is the uncertainty within the measured salt intrusion length. It was shown in a sensitivity analysis that this parameter has a large influence on the predicted discharge: a 10% decrease/increase in L results in a 25% increase/decrease in $Q_{f,}$ implying an elasticity of 2.5. Therefore it is very important to work adequately when performing field work.

5.4. Conclusion

It can be concluded that there appears to be potential for mutual model improvement when linking a rainfall-runoff model to a steady state salt intrusion model. The prediction of a salt intrusion model could be improved by a more realistic input value for the fresh water discharge, originating from the prediction of a rainfall-runoff model. Meanwhile, the response of the upscaled rainfall-runoff model can be validated by the back calculated fresh water discharge of the salt intrusion model.

The single discharge prediction that was compared showed good resemblance. Therefore hypothesis 3 is confirmed: Linking a rainfall-runoff model to a steady state salt intrusion model offers mutual model improvement. However, this confirmation is only based on preliminary research, to really verify the mutual model improvement, it is recommended to perform extra research, whereby a longer time series is investigated. Future application could be the linkage to a non-steady state salinity model. Such a model requires continuous input of discharge data, which could be supplied by the rainfall-runoff model.

To improve the method of upscaling it is recommended to use a distributed precipitation forcing to account for the heterogeneity in rainfall and apply more advanced methods to account for for possible delay due to channel routing and wave attenuation.

The relative good performance of the upscaled model when validated on the salt intrusion model confirms the predictive power of the model. Since the model was calibrated on stage, this approach is implicitly validated as well. This adds to the credibility of the proposed method of calibration on stage.

CHAPTER 6

Concluding remarks

In this research it was shown that calibrating a rainfall-runoff model on stage reduces discharge uncertainty in scarcely gauged basins; topography driven rainfall-runoff models are better transferable than lumped rainfall-runoff models; and linking a salt intrusion model to a rainfall-runoff model offers mutual model improvement.

Calibration on stage gives plausible results in the poorly gauged Endau River Catchment and has potential for being applied more widely. It offers a simple though robust method that eliminates the requirement of error-prone discharge measurements. It was found that in the method whereby the rating curve is defined by two parameters, the rating curve parameters are well defined and optimise to values that correspond to the physical property they represent. This implies that the addition of rating curve parameters does not increase the risk of over-parameterisation. In a sensitivity analysis, it was found that the reanalysed rating curves are insensitive to model structure and not very sensitive to the potential evaporation while a strong sensitivity was found to the precipitation forcing. The models calibrated on stage are well able to mimic the reanalysed hydrographs, while it was impossible to find a good fit for the original 'observed' hydrographs. During validation on a different period and a different catchment the models calibrated on stage were able to simulate the hydrologic response as well.

The transfer of topography driven models to a different catchment gives better results than the transfer of lumped models. This implies that the process realism and predictive power of the topographical models is higher than of the lumped models. This finding is also important for studies on ungauged basins: a calibrated topographical model found for a gauged catchment can be readily applied to a similar ungauged catchment in a comparable climate.

Finally, there appears to be potential for mutual model improvement when linking a rainfall-runoff model to a steady state salt intrusion model. This method could improve the prediction of the salt intrusion length by the salt intrusion model and function as a way of validation for the rainfall-runoff model. Future applications could be the linkage to a dynamic salt intrusion model. This model requires continuous input of discharge data, which could be supplied by the rainfall-runoff model.

Appendix A – Catchment delineation

This appendix elaborates on the method to delineate the various catchments that were studied in this research. The basis for the catchment boundaries is a Digital Elevation Model (DEM) that was obtained from radar altimetry in the Shuttle Radar Topography Mission (SRTM). The DEM data from this space mission covers most of the populated regions of the world and is freely available at a spatial resolution of 3 x 3 arc seconds, equal to about 90 x 90 meters around the equator. The DEM data has a vertical accuracy of 16 meters and a horizontal accuracy of 20 meters (Berry *et al*., 2007). To obtain actual catchment boundaries, the DEM data was further processed in the GIS program ArcGIS 10. The DEM was first corrected for possible ‘pit’ cells: cells that have an unrealistically low value, causing wrong drainage directions when not corrected for. After correction of these pits, the slopes between neighbouring cells were calculated. From this slope model, the drainage direction was calculated, whereby was assumed that water will flow in the direction with the steepest gradient, according to the D8 method, introduced by O’Callaghan *et al.* (1984). Subsequently, one can determine the number of cells that flow into a certain cell (pour point cell). Each pour point has its own specific catchment area. Now it is just a matter of determining a pouring point, which functions as outflow of the to-be-determined catchment.

Appendix B – Field survey at station 2 in sub-catchment 2

At gauging station 2 in sub-catchment 2, the meteorological station was inspected (Figure 23) and a cross-section survey was performed together with discharge measurements. The cross-section was measured using a diver attached to a couple of bricks. This construction was lowered into the river from the bridge (visible in Figure 22). Subsequently, every meter a measurement was taken, whereby the brick-diver construction was placed on the bottom for a few seconds, so that it measured the water pressure and the air pressure. Subsequently, the brick-diver construction was lifted out of the water, so that it measured only the air pressure, and so on. In this way the individual measuring points could easily be distinguished. The result is shown in Figure 21. From this cross-section the side slopes (m_1 and m_2) have been estimated by taking a linear regression. It was assumed that this regression line can be extrapolated when the water table rises. This assumption was validated by a field survey, whereby it was observed that even during the highest recorded water level, the assumed cross-section is still a good approximation.

The discharge was determined using the velocity-area method (Luxemburg *et al.*, 2011) on two different days. For the first day, the area was measured as described above, while for a second day only the height difference in the middle was measured, from which the corresponding area was calculated using the height-area relation. Therefore, the measurement on the first day is considered more reliable than the measurement on the second day. The velocity at the surface was measured at three points over the width of the river. At each point five measurements were taken; the velocity at the surface was calculated as the average of these five measurements. A float made of a plastic 1.5 l bottle was half filled with sand, attached to a rope and lowered into the river. Subsequently, the float was transported by the river over a 10 meter distance while the time was measured. The velocity was calculated as $v = s/t$. This velocity was corrected by a value 0.9 to obtain the mean velocity in the vertical (Luxemburg *et al.*, 2011).

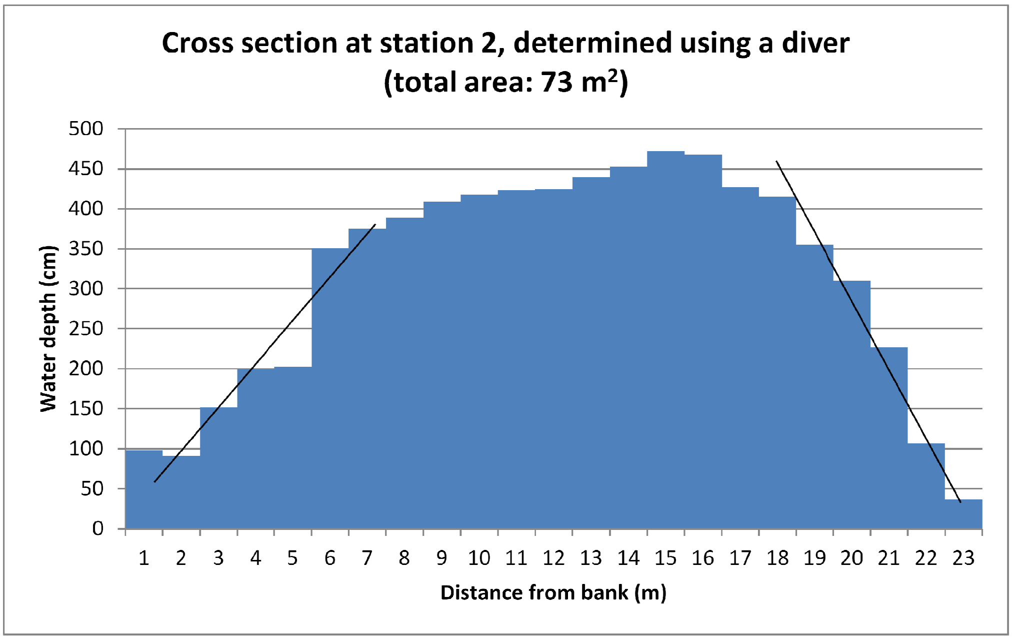


Figure 21 - Cross-section measured from bridge at station 2 in sub-catchment 2. The black lines represent the regression from which the side slopes have been determined.

Figure 22 – An impression of the gauging station of sub-catchment 2. The water flows in the direction of the bridge.

Figure 23 – An impression of the meteorological station of sub-catchment 2. It can be seen that obstacles are present nearer than five times the obstacle height of the measuring equipment, due to which measurements may be inaccurate.

Appendix C – Double mass analyses precipitation

Double mass curves of the satellite data and the filled up ground data are shown in Figure 24 and Figure 25. On the x-axis the accumulated values of the data series under investigation are plotted, on the y-axis the average of all other accumulated data series are plotted. The period that is shown ranges from 2003 to 2013.

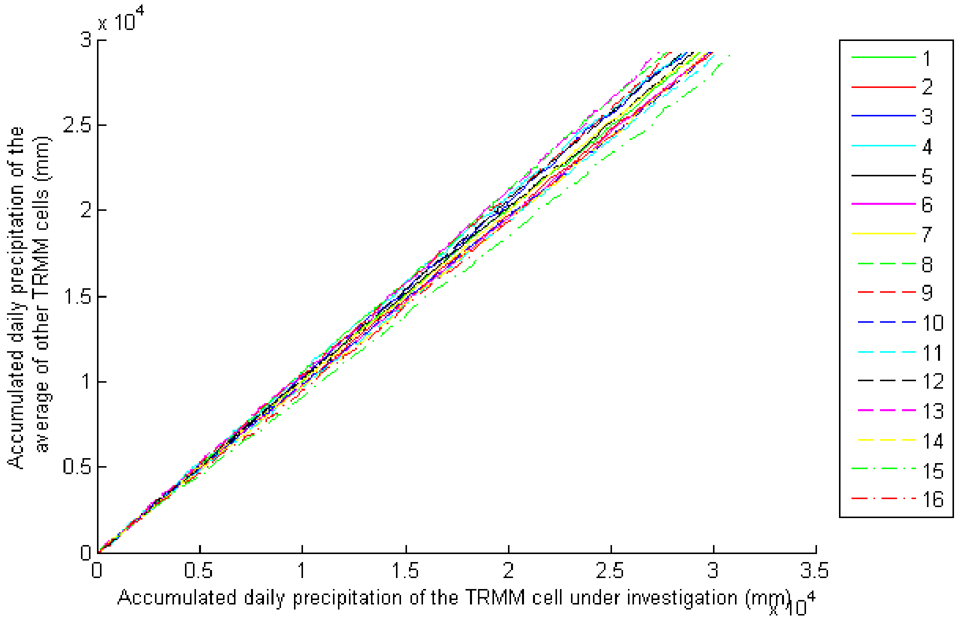


Figure 24 – Double mass curves of the TRMM data, whereby the accumulated values of the TRMM cell under investigation are plotted against accumulated values of the average of other TRMM cells over the same period.

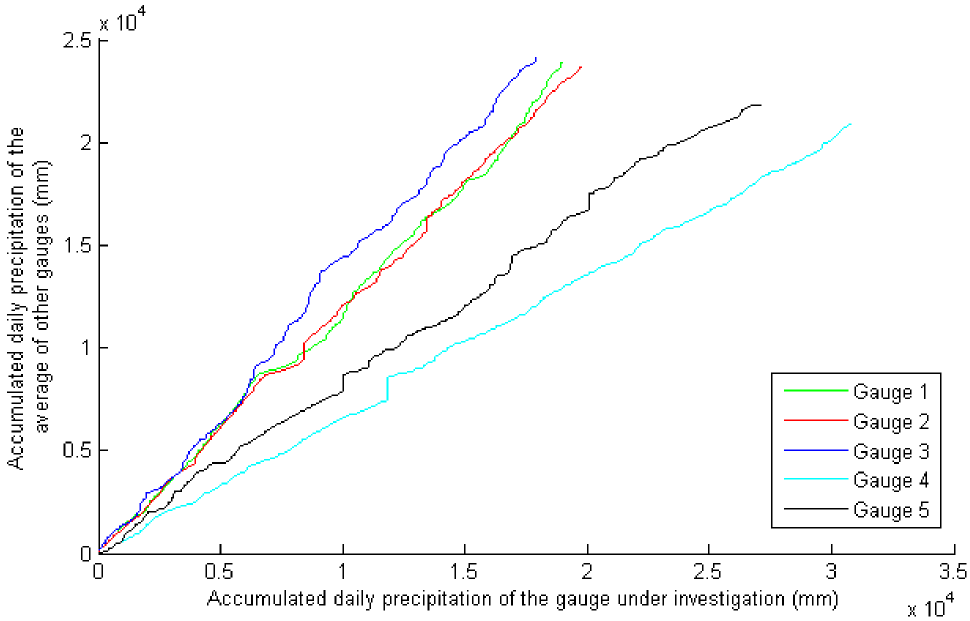


Figure 25 – Double mass curves of the ground data, whereby the accumulated values of the station under investigation are plotted against accumulated values of the average of other stations over the same period.

Appendix D – Derivation of K_s

As in many other cases, the receding limb of the original and reanalysed hydrographs of the Endau River catchment can be described by a linear storage-discharge relationship and therefore plots as a straight line after a log transform of the discharge, whereby the offset is determined by the initial storage condition. The slope of the straight line represents the depletion coefficient, the time scale of the outflow process, K_s, by: $K_s = -1/slope$, this relation is analytically derived in Eq. 44 to 53 (Tallaksen, 1995).

(44) $Q = \frac{1}{k} S$

In a dry period the inflow to the slow reservoir is 0, so the water balance is:

(45) $\frac{ds}{dt} = -Q$

Combination and manipulation of the above two equations gives:

(46) $\frac{dS}{dt} = -\frac{1}{k} S$

(47) $\frac{1}{S} dS = -\frac{1}{k} dt$

(48) $\int \frac{1}{S} dS = \int -\frac{1}{k} dt$

(49) $ln\, S = -\frac{1}{k} t + C$

(50) $S = e^{-\frac{1}{k}t + C} = Ce^{-\frac{1}{k}t}$

(51) $S = S_0\ at\ t = 0$ so $S = S_0\, e^{-\frac{1}{k}t}$

Substitution of Eq. 44 gives:

(52) $Qk = S_0\, e^{-\frac{1}{k}t}$ so $Q = \frac{S_0}{k}\, e^{-\frac{1}{k}t} = Q_0 e^{-\frac{1}{k}t}$

(53) $log\, Q = log(Q_0 e^{-\frac{1}{k}t}) = log\, Q_0 + log(e^{-\frac{1}{k}t}) = log\, Q_0 - \frac{1}{k} t\, log\, e = -\frac{1}{k} t + log\, Q_0$

All reanalysed hydrographs were analysed in this way to find their characteristic K_s value. First, the receding limbs of the considered hydrograph were manually selected using only recession periods longer than the selected threshold of 6 days (Figure 26). Subsequently, all receding limbs found were plotted together in one figure to check their resemblance (Figure 27). The final value for K_s was set as the average of all values. It was checked with automatic calibration whether the model optimises to the same value, this was the case.

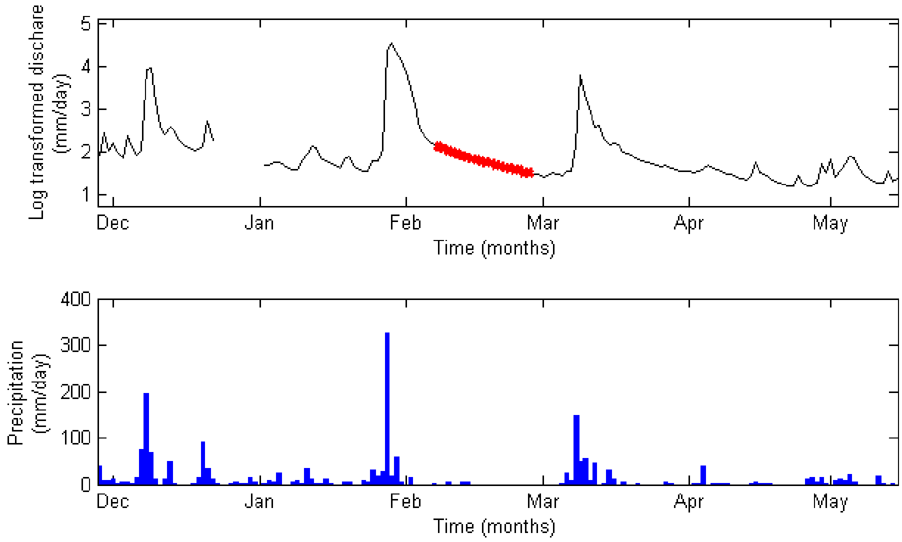


Figure 26 - Manual selection of receding limb during dry conditions.

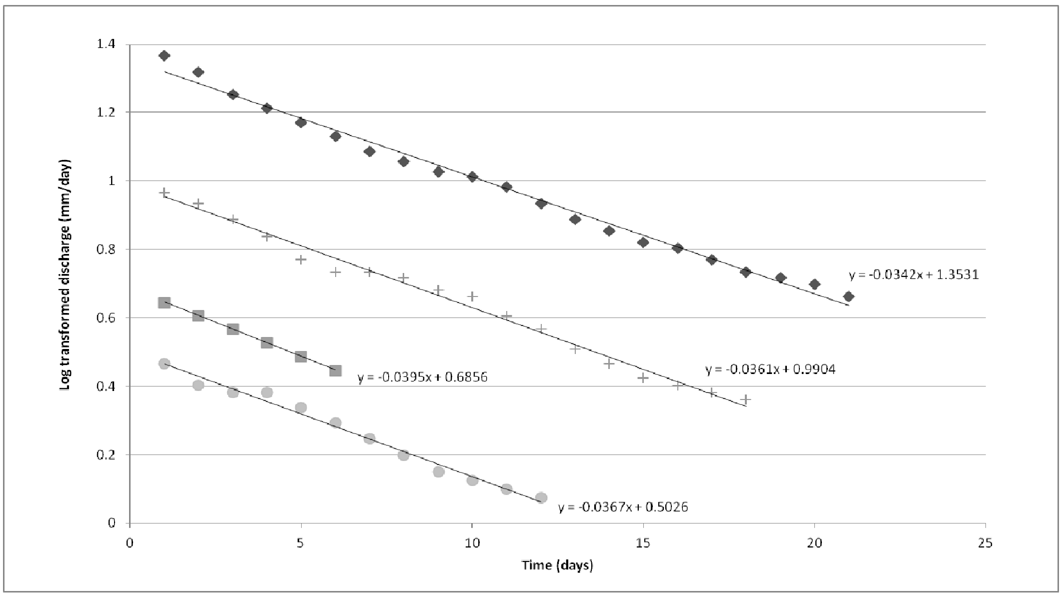


Figure 27 - Selected receding limbs with linear regression lines and corresponding equations. For this hydrograph K_s was set at 27 days. NB this is figure shows the receding limbs of a model that has not been used in the remainder of this report.

Appendix E – Prior parameter ranges

Table 7 - Prior parameter ranges of the lumped model, the topography driven model and the rating curve model.

Lumped model

Parameter	I_{max}	p	$S_{u,max}$	β	P_{max}	K_f	K_s	D	C_{max}	K_0	$S_{f,max}$
Unit	mm	-	mm	-	mm/d	d	d	-	mm/d	d	mm
Lower limit	1	0.5	50	0.1	0.01	1	10	0.1	0	1	10
Upper limit	8	0.5	700	5	5	10	500	1	3	1	200

Topograpy driven model

Parameter	F_H	β_H	$S_{u,max,H}$	p_H	D_H	$K_{f,H}$	F_W	β_W	$S_{u,max,W}$	$K_{f,W}$	K_s	C_{max}
Unit	mm/d	-	mm	-	-	d	mm/d	-	mm	d	d	mm/d
Lower limit	1	0.1	50	0.5	0.1	1	0	0.1	20	1	5	0.01
Upper limit	8	5	700	0.5	1	20	5	5	500	10	200	2

Rating curve model

Parameter	a	b	h_{0-1}	h_{0-2}	slope
Unit	m^{3-b}/s	-	m	m	-
Lower limit	free	1	27.5	8	0
Upper limit	free	3	32	12	0.002

Appendix F – Posterior parameter ranges

Table 8 – Posterior model parameters and rating curve parameters of the six tested models.

CATCHMENT 1

Lumped model

Parameter	I_{max}	p	$S_{u,max}$	β	P_{max}	K_f	K_s	D	C_{max}	K_0	$S_{f,max}$
Unit	mm	-	mm	-	mm/d	d	d	-	mm/d	d	mm
Lower limit	3,1	0,50	327	2,4	2,6	2,7	40	0,70	0,74	1,0	107
Upper limit	4,1	0,50	336	2,6	2,6	3,0	40	0,70	1,1	1,0	111

Rating curve model

Parameter	a	b	h_{0-1}
Unit	m^{3-b}/s	-	m
Lower limit	17,3	1,38	28,8
Upper limit	17,3	1,38	28,8

Topography driven model

Parameter	F_H	β_H	$S_{u,max,H}$	p_H	D_H	$K_{f,H}$	F_W	β_W	$S_{u,max,W}$	$K_{f,W}$	K_s	C_{max}
Unit	mm/d	-	mm	-	-	d	mm/d	-	mm	d	d	mm/d
Lower limit	5,3	2,0	430	0,50	0,27	1,4	2,4	0,82	133	3,0	37	0,041
Upper limit	6,8	2,5	463	0,50	0,31	1,5	3,1	0,92	177	3,1	37	0,10

Rating curve model

Parameter	a	b	h_{0-1}
Unit	m^{3-b}/s	-	m
Lower limit	16,8	1,47	28,9
Upper limit	16,8	1,47	28,9

CATCHMENT 2

Lumped model

Parameter	I_{max}	p	$S_{u,max}$	β	P_{max}	K_f	K_s	D	C_{max}	K_0	$S_{f,max}$
Unit	mm	-	mm	-	mm/d	d	d	-	mm/d	d	mm
Lower limit	2,0	0,50	77,0	2,5	0,41	4,7	40,0	0,67	0,70	1,0	138
Upper limit	2,6	0,50	84,1	2,9	0,48	4,8	40,0	0,66	0,81	1,0	176

Rating curve model

Parameter	a	b	h_{0-2}
Unit	m^{3-b}/s	-	m
Lower limit	9,34	1,45	11,3
Upper limit	9,34	1,45	11,3

Lumped model

Parameter	I_{max}	p	$S_{u,max}$	β	P_{max}	K_f	K_s	D	C_{max}	K_0	$S_{f,max}$
Unit	mm	-	mm	-	mm/d	d	d	-	mm/d	d	mm
Lower limit	2,7	0,50	71,1	2,8	0,38	4,2	42	0,83	0,53	1,0	176
Upper limit	4,0	0,50	75,2	2,9	0,65	4,5	42	0,76	0,60	1,0	186

Rating curve model

Parameter	slope	h_{0-2}
Unit	-	m
Lower limit	0,000310	11,3
Upper limit	0,000310	11,3

Topography driven model

Parameter	F_H	β_H	$S_{u,max,H}$	p_H	D_H	$K_{f,H}$	F_W	β_W	$S_{u,max,W}$	$K_{f,W}$	K_s	C_{max}
Unit	mm/d	-	mm	-	-	d	mm/d	-	mm	d	d	mm/d
Lower limit	3,0	3,0	399	0,50	0,21	3,7	0,41	0,66	93	5,9	35	0,04
Upper limit	5,5	3,8	495	0,50	0,24	4,6	0,68	0,76	104	6,0	35	0,10

Rating curve model

Parameter	a	b	h_{0-2}
Unit	m^{3-b}/s	-	m
Lower limit	9,70	1,40	11,4
Upper limit	9,70	1,40	11,4

Topography driven model

Parameter	F_H	β_H	$S_{u,max,H}$	p_H	D_H	$K_{f,H}$	F_W	β_W	$S_{u,max,W}$	$K_{f,W}$	K_s	C_{max}
Unit	mm/d	-	mm	-	-	d	mm/d	-	mm	d	d	mm/d
Lower limit	3,4	3,3	423	0,50	0,24	3,0	1,1	0,79	80	4,7	35	0,15
Upper limit	4,5	3,7	528	0,50	0,37	4,2	2,5	0,95	87	5,0	35	0,23

Rating curve model

Parameter	slope	h_{0-2}
Unit	-	m
Lower limit	0,000320	11,4
Upper limit	0,000320	11,4

Appendix G – Calibration and validation objective scores

Table 9 shows the average calibration and validation objective scores of the behavioural models, where γ is Euclidean distance (whereby a lower value indicates a better score) and A to E are objective functions (whereby a higher value represents a better score).

Objectives A to E denote:

A. Nash-Sutcliffe Efficiency of the flow duration of the log of the flows (NSE_{logFDC})
B. Nash-Sutcliffe Efficiency of the flow duration curve (NSE_{FDC})
C. Nash-Sutcliffe Efficiency of the log of flows (NSE_{logQ})
D. Nash-Sutcliffe Efficiency of the flows (NSE_{Q})
E. Autocorrelation (AC)

Note that the Euclidean distance is given for the calibration period. This was only done to be able to compare the scores with the scores during validation, but was not used during the calibration process.

Table 9 - Objective scores during calibration and validation. NB C1, 3 par denotes: sub-catchment 1, 3-parameter definition.

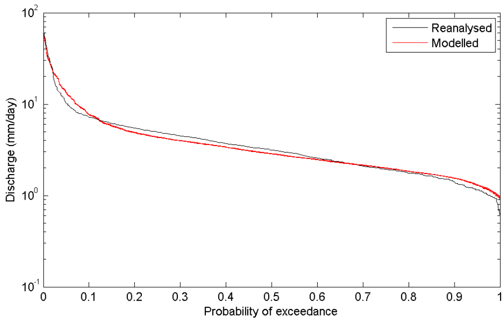
Model type		Calibration						Split-record validation						Different catchment validation					
		A	B	C	D	E	γ	A	B	C	D	E	γ	Transferred to	A	B	C	D	γ
Lumped	C1, 3 par	0,98	0,97	0,58	0,60	0,97	0,58	0,55	-0,20	0,08	-1,1	0,78	2,6	C2, 3 par	0,83	0,92	0,52	0,38	0,81
														C2, 2 par	0,77	0,87	0,50	0,40	0,83
Topo	C1, 3 par	0,98	0,95	0,43	0,62	0,97	0,69	0,46	-0,28	-0,02	-1,1	0,75	2,7	C2, 3 par	0,96	0,92	0,40	0,09	1,1
														C2, 2 par	0,97	0,99	0,55	0,28	0,84
Lumped	C2, 3 par	0,99	0,99	0,61	0,52	0,80	0,65	0,97	0,78	0,43	0,27	0,88	0,96	C1, 3 par	0,83	0,89	0,11	0,45	1,1
	C2, 2 par	0,99	0,99	0,61	0,51	0,75	0,68	0,98	0,83	0,44	0,32	0,87	0,90	C1, 3 par	0,62	0,76	-0,24	0,24	1,5
Topo	C2, 3 par	0,99	0,99	0,62	0,54	0,96	0,59	0,97	0,90	0,42	0,41	0,99	0,84	C1, 3 par	0,97	0,81	0,47	0,57	0,71
	C2, 2 par	0,98	0,98	0,63	0,52	0,97	0,60	0,98	0,91	0,48	0,45	0,99	0,76	C1, 3 par	0,97	0,87	0,40	0,60	0,73

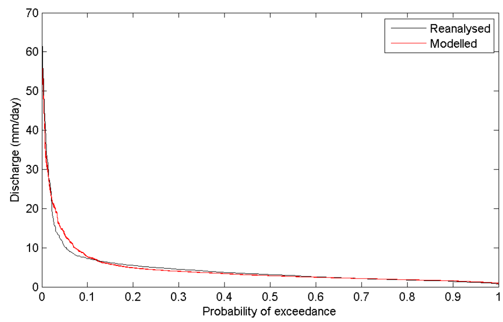
Appendix H – Hydrographs and flow duration curves during calibration

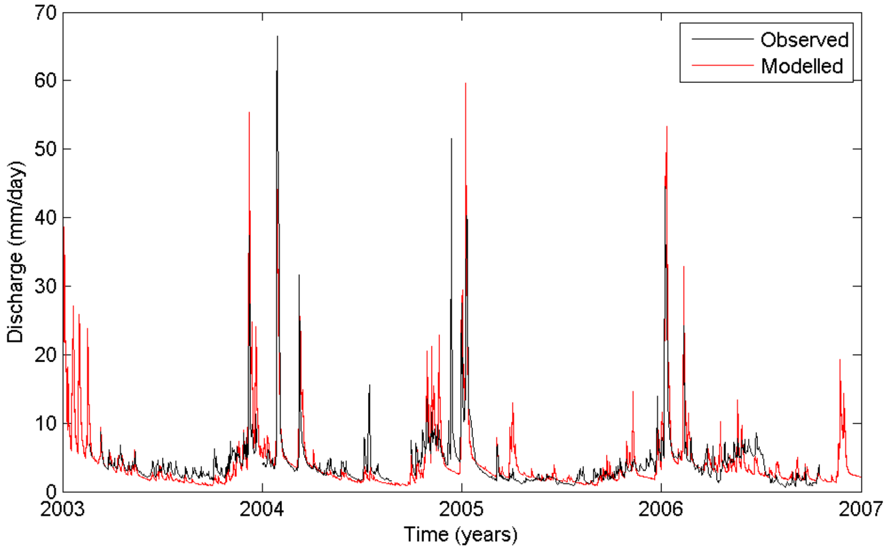
The figures below show the flow duration curves and hydrographs of the behavioural parameter sets found during calibration over the period 2003 – 2007. The year 2003 – 2004 was taken as spin up time and was not accounted for in calculating the objectives or construction of the flow duration curves.

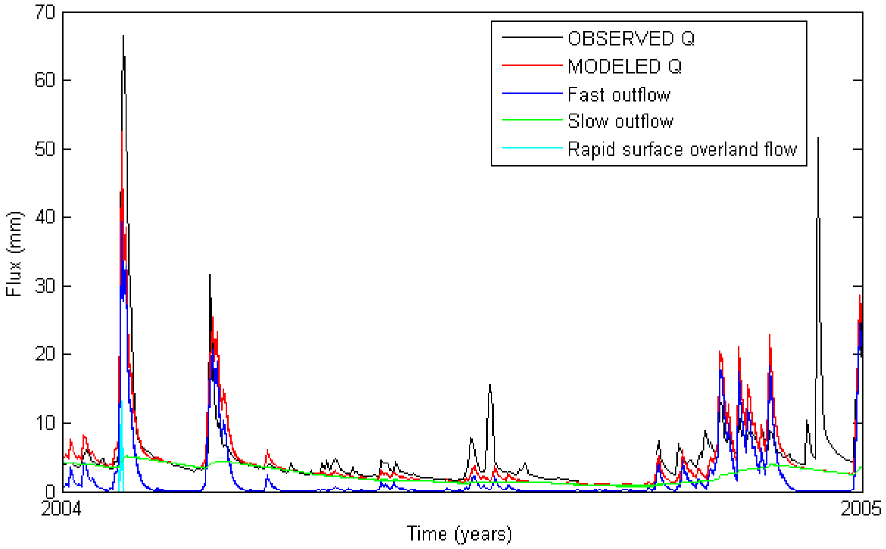
The order of the figures is as follows: the first two figures describe the flow duration curve of the log of the flows and the flow duration curve of the flows; the third figure shows the hydrograph over the entire period 2003 – 2007; the fourth and last figure shows the decomposed hydrograph, implying that the hydrograph is split up into its various flow components. For the lumped model, the components that are indicated are the fast outflow (Q_f), the slow outflow (Q_s) and the rapid surface overland flow (Q_0). In the topography driven model we distinguish between the runoff from the wetlands (Q_{fW}), the runoff from the hillslopes (Q_H) and the slow outflow from the groundwater reservoir (Q_s).

Sub-catchment 1, Lumped

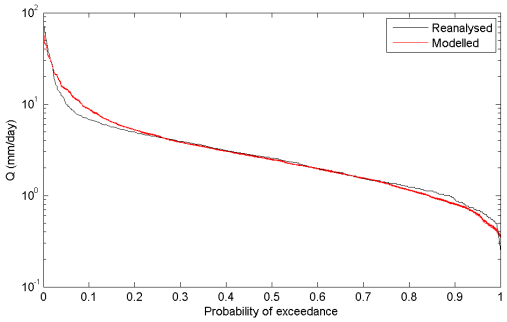


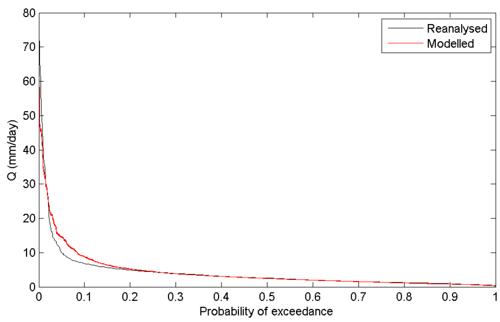


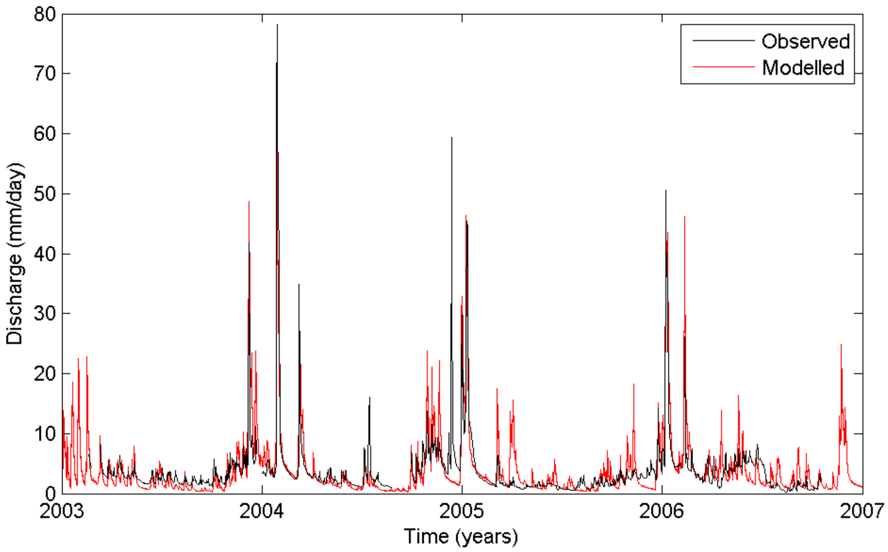


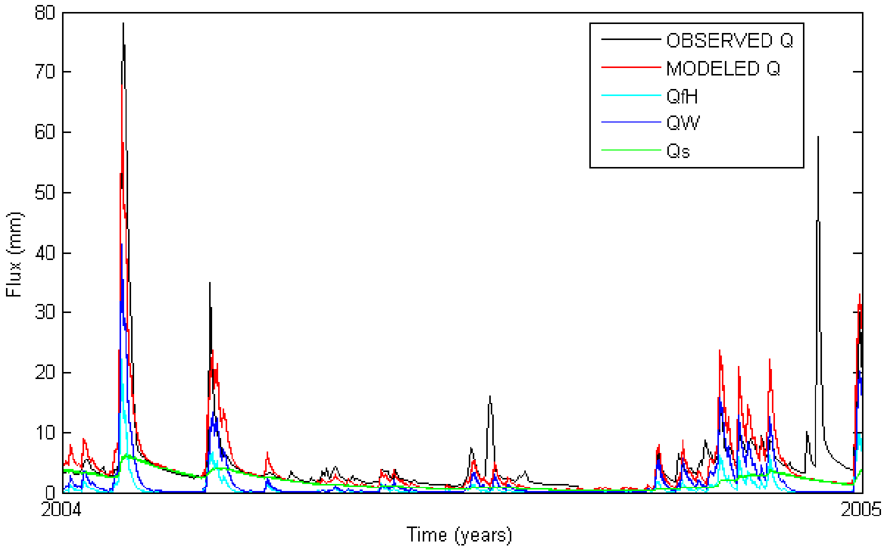


Sub-catchment 1, Topography driven

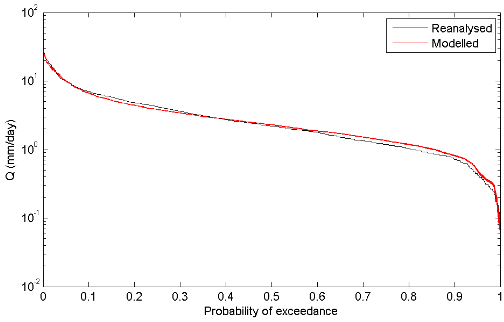


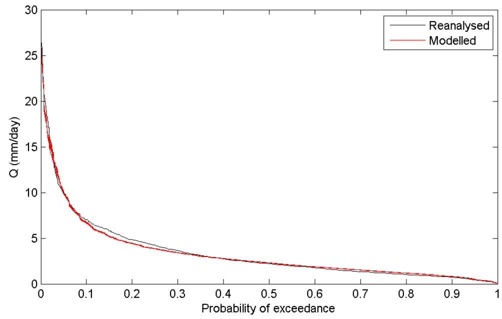


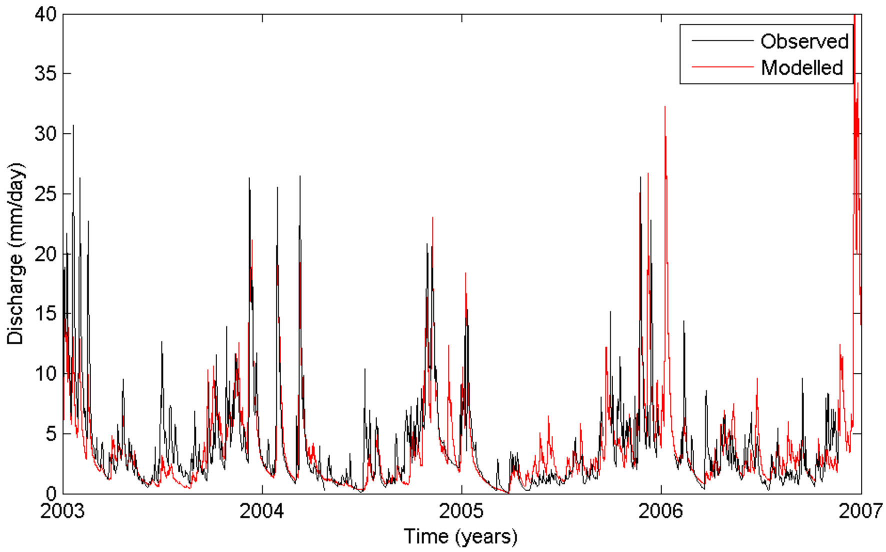


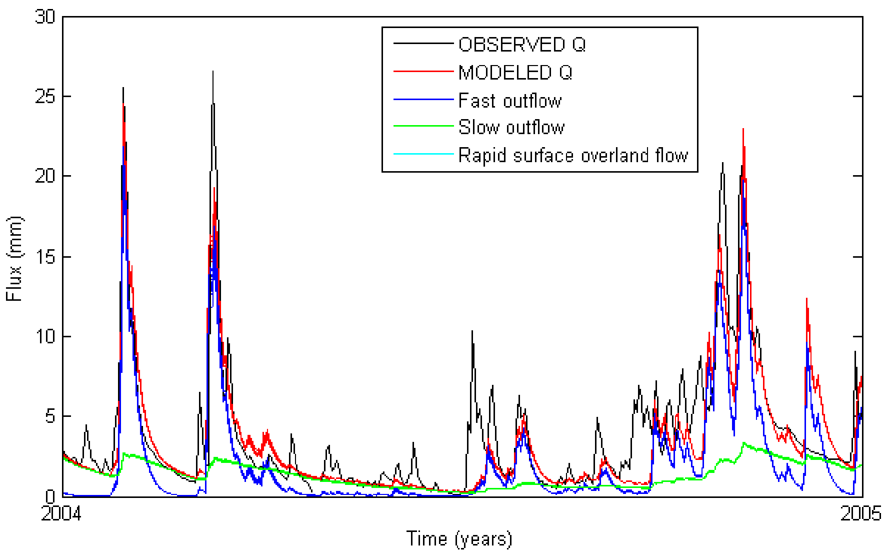


Sub-catchment 2, Lumped, 3 parameters

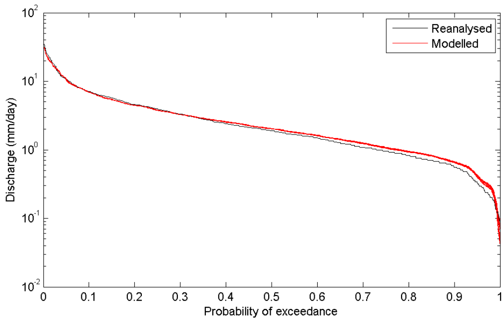


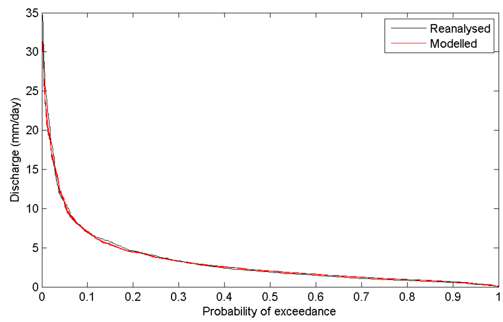


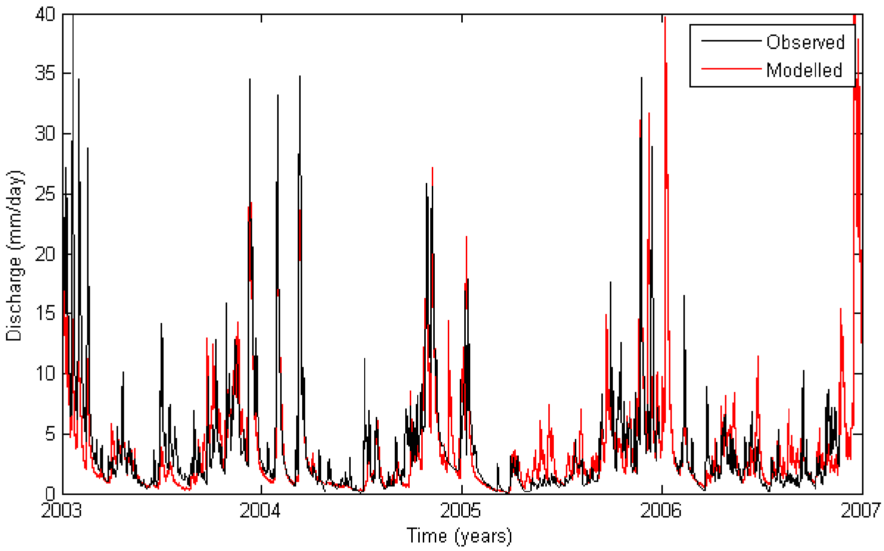


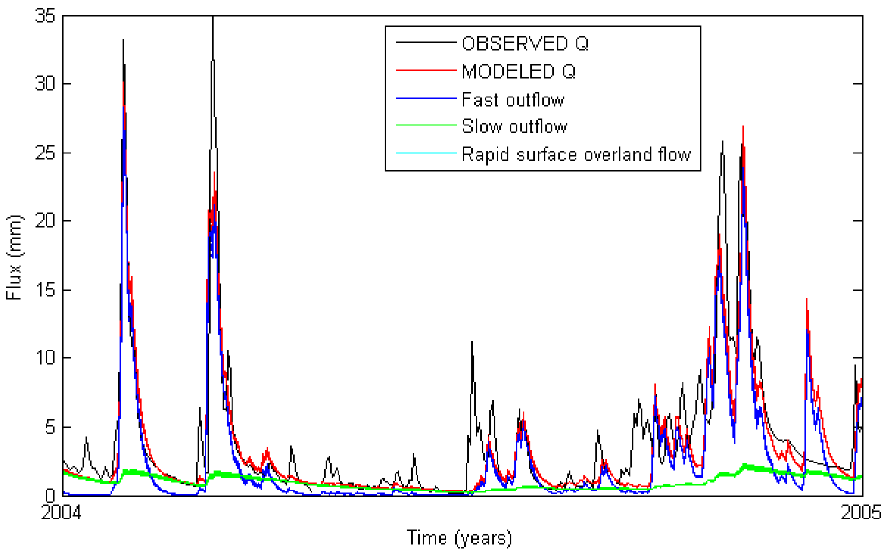


Sub-catchment 2, Lumped, 2 parameters

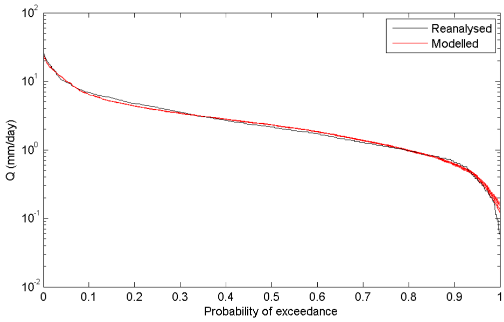


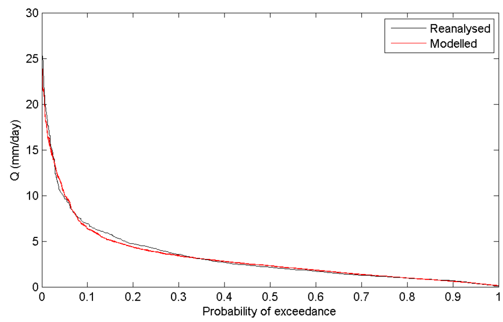


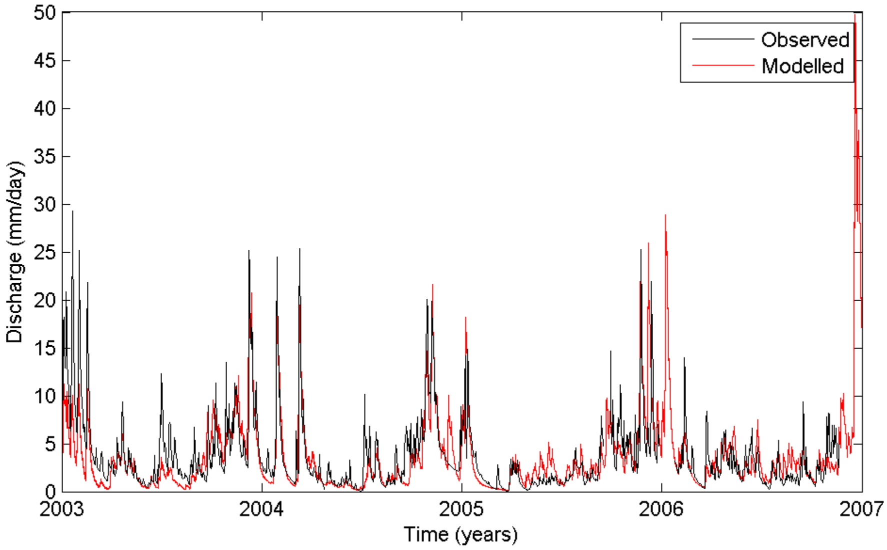


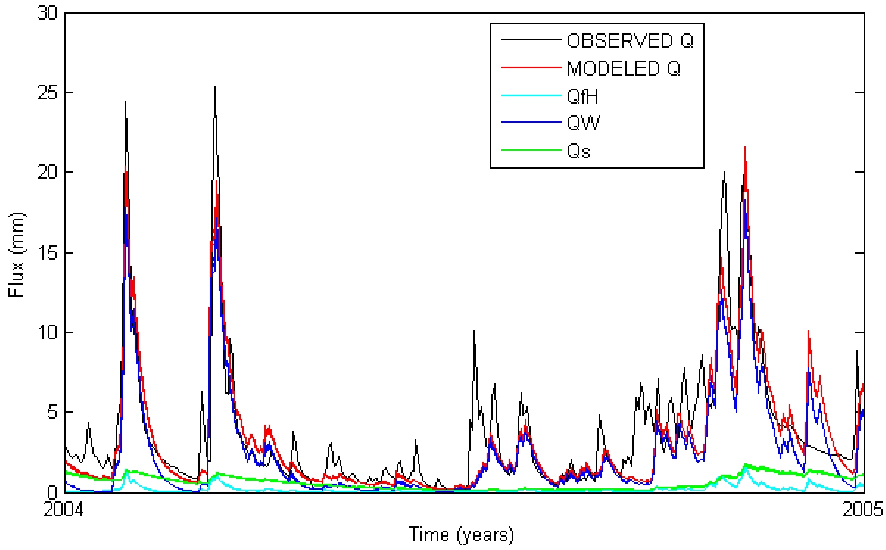


Sub-catchment 2, Topography driven, 3 parameters

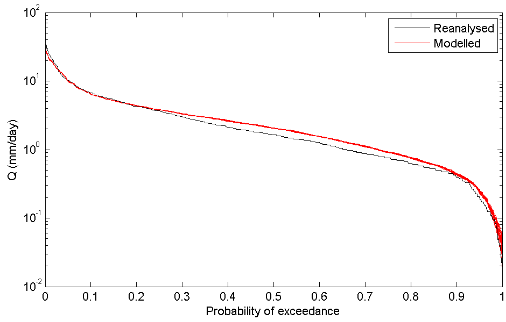


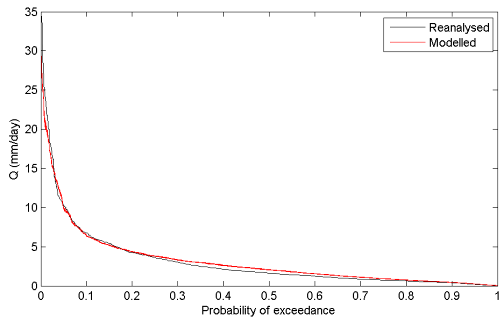


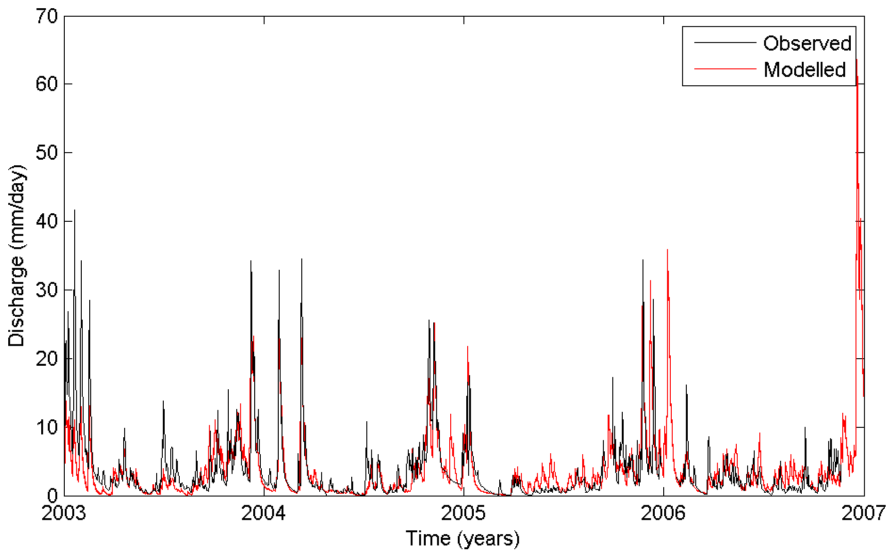


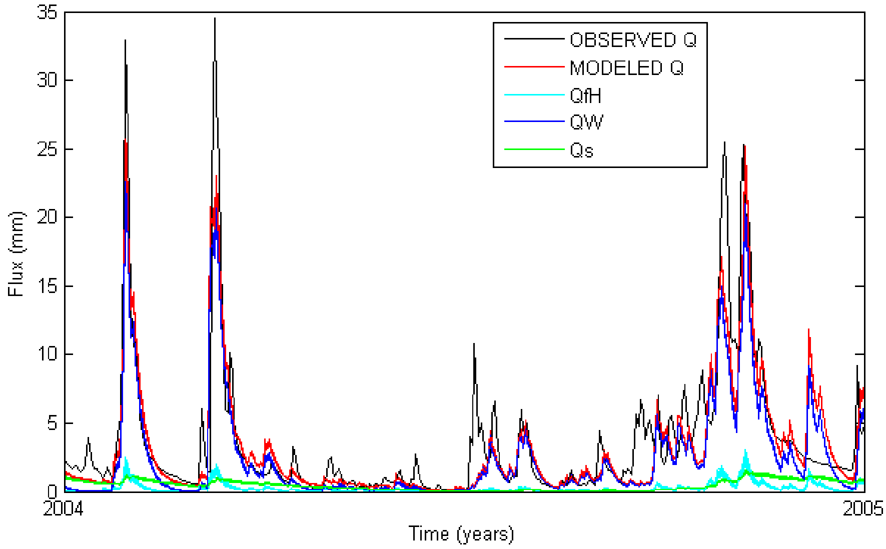


Sub-catchment 2, Topography driven, 2 parameters





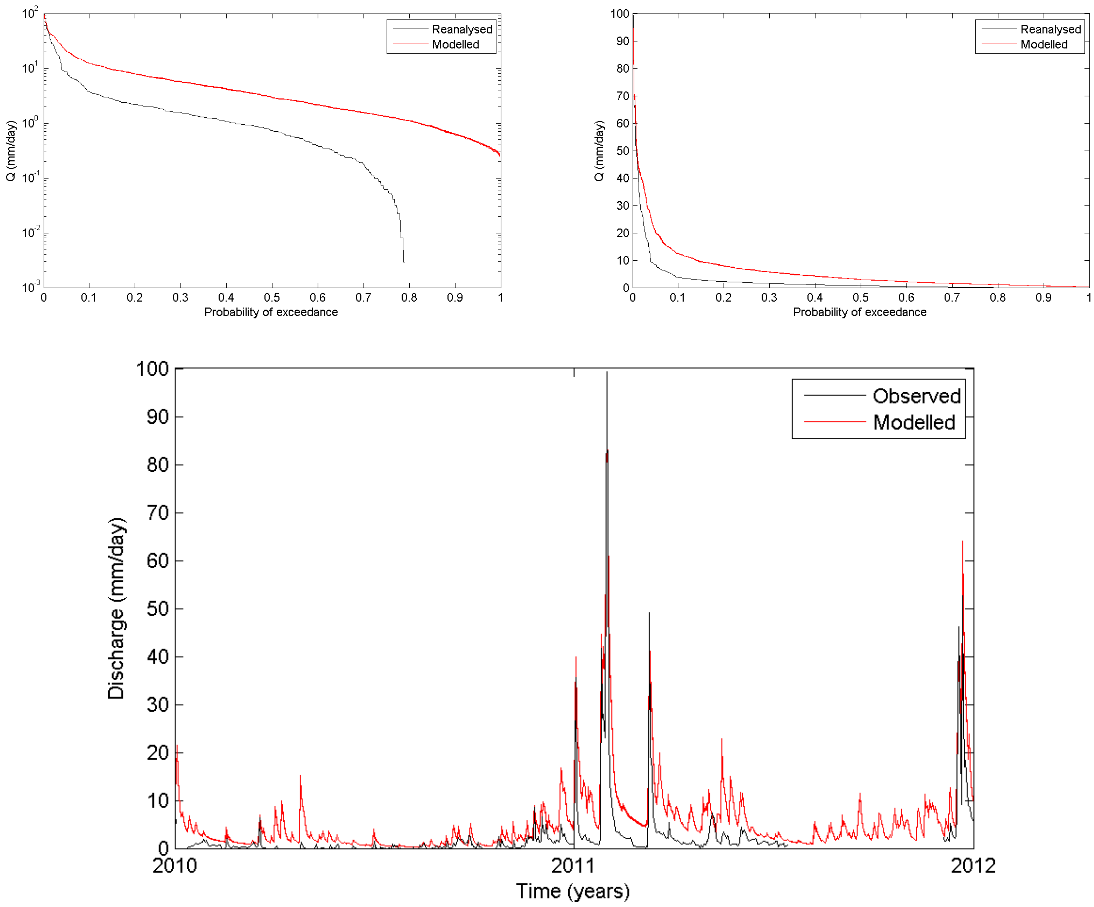




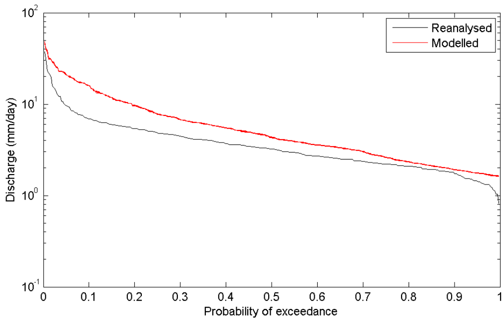
Appendix I – Hydrographs and flow duration curves during validation on a different period

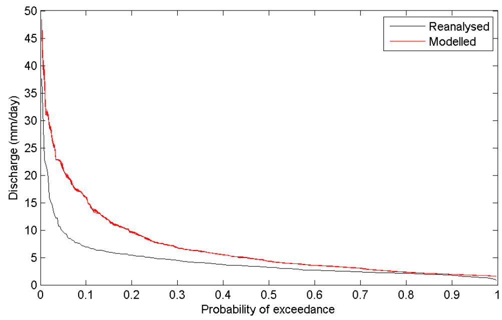
The figures below show the flow duration curves and hydrographs during split-record validation. Sub-catchment 1 was validated over the years 2007 – 2009, whereby 2007 – 2008 was used as spin up time and therefore not shown in the figures. Sub-catchment 2 was validated over the period 2009 – 2012, with a spin-up time of one year which was not shown in the figures. To illustrate the bad performance of the models of sub-catchment 1 in the period 2009 – 2012, these figures are shown as well.
The order of the figures is as follows: the first two figures describe the flow duration curve of the log of the flows and the flow duration curve of the flows; the third figure shows the hydrograph.

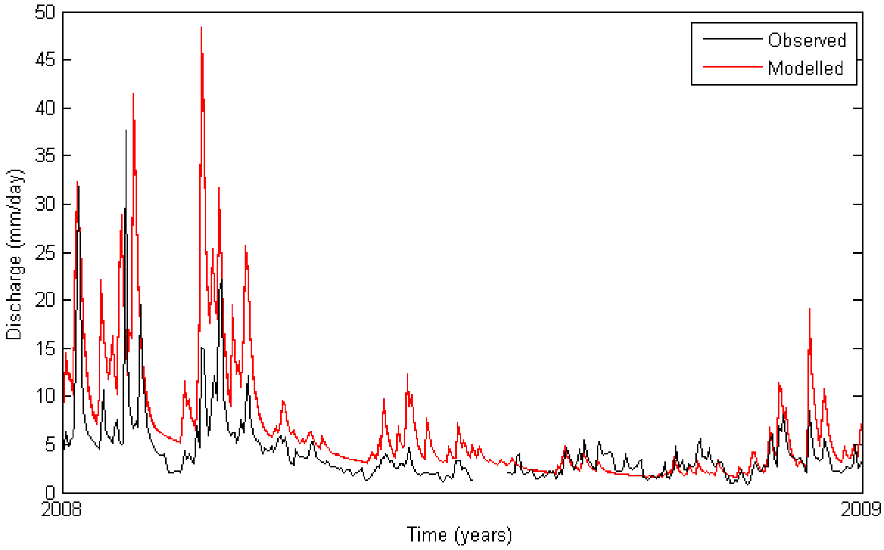
Sub-catchment 1, Topography driven (period 2009 – 2012)



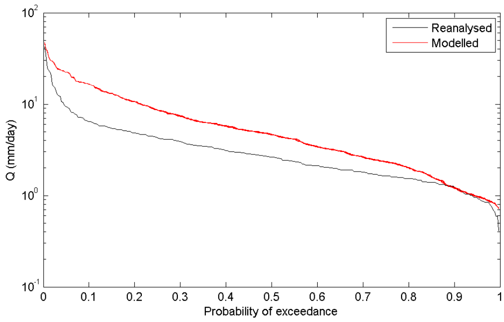
Sub-catchment 1, Lumped

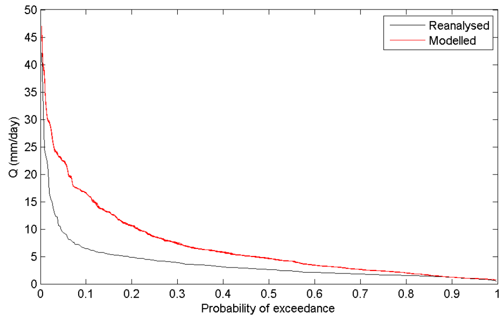


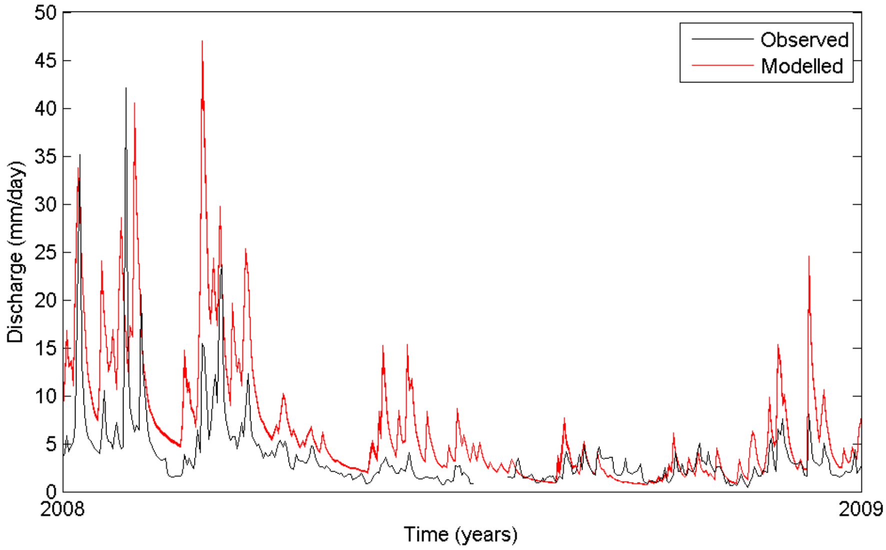




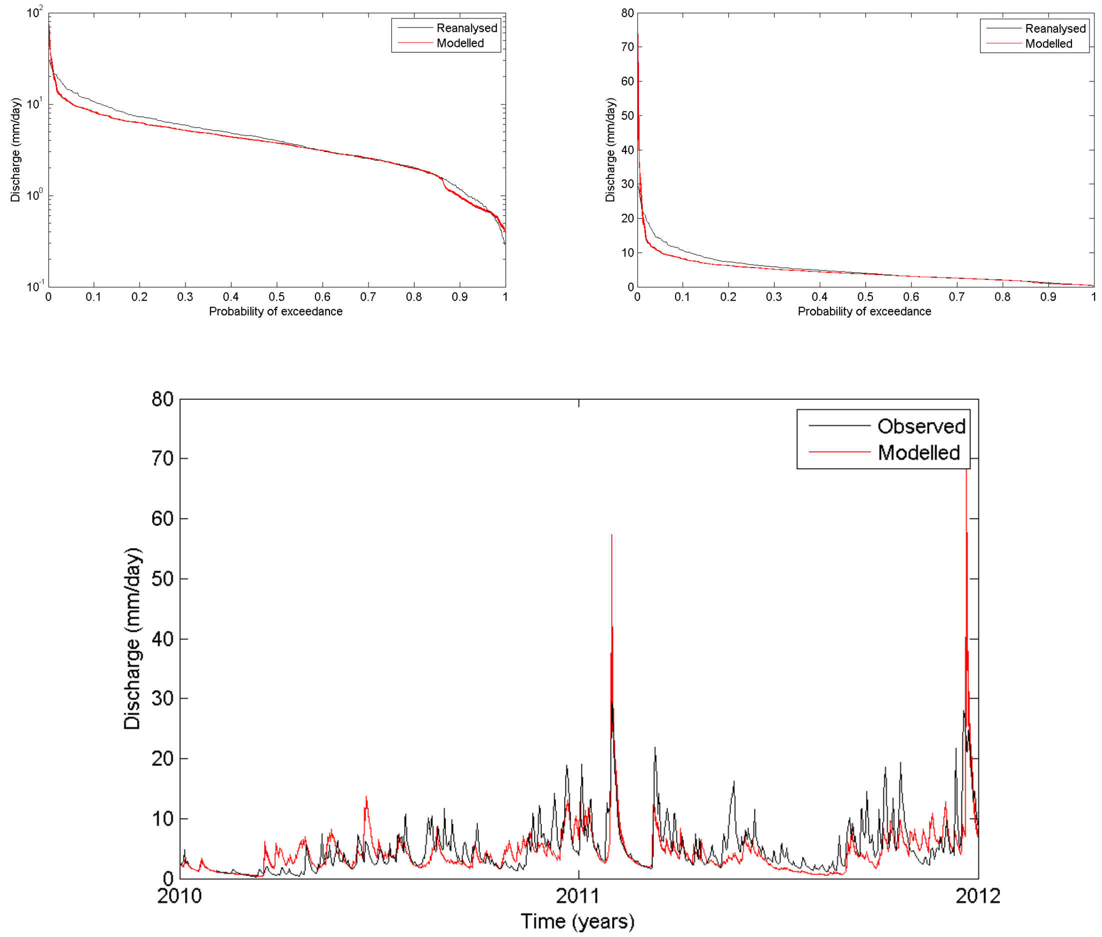
Sub-catchment 1, Topography driven



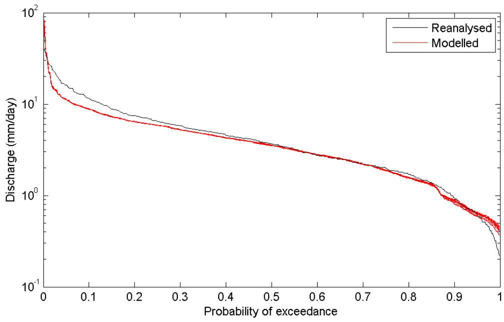


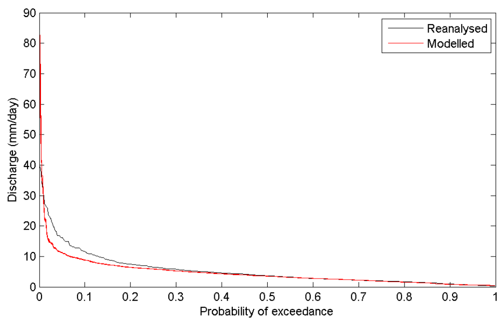


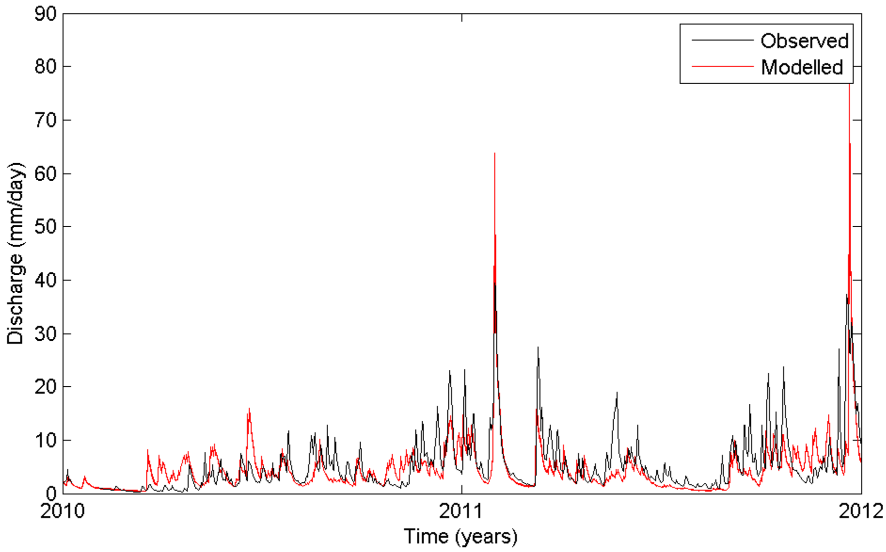
Sub-catchment 2, Lumped, 3 parameters



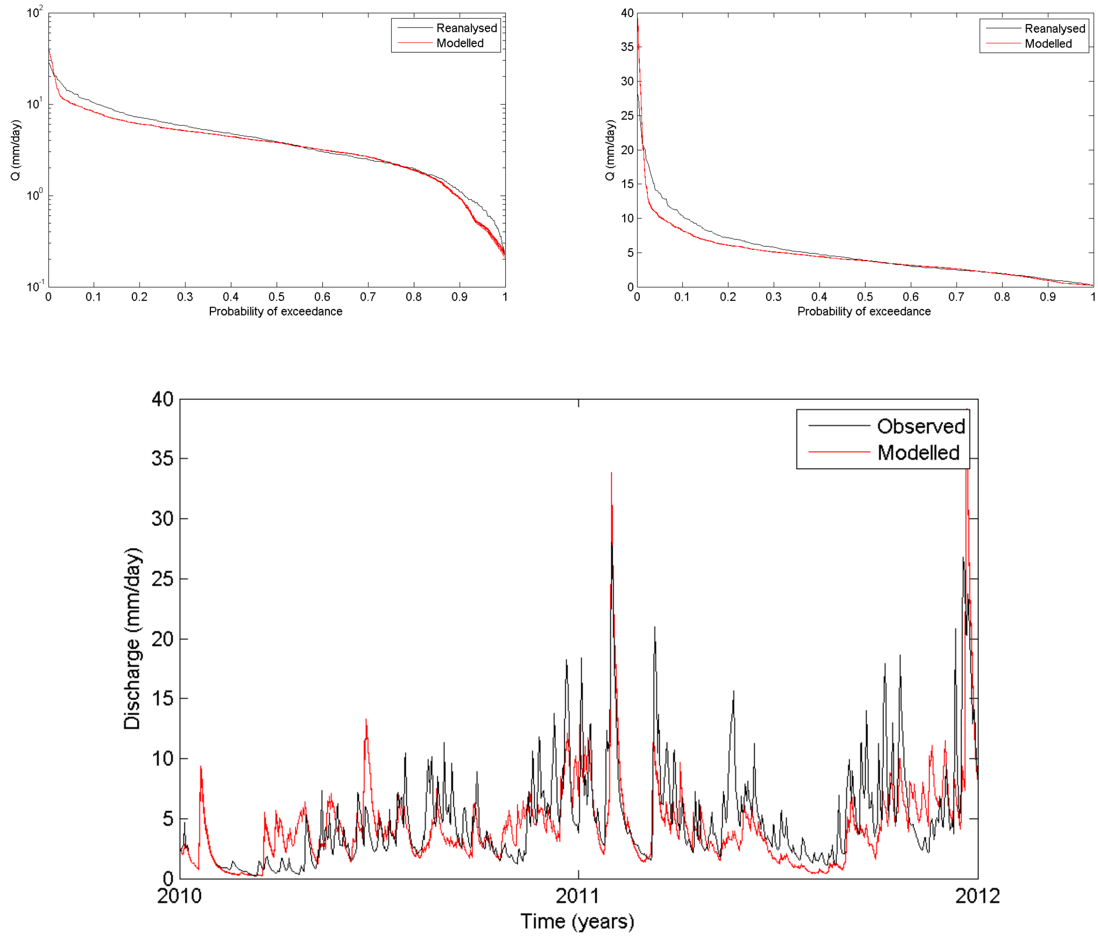
Sub-catchment 2, Lumped, 2 parameters



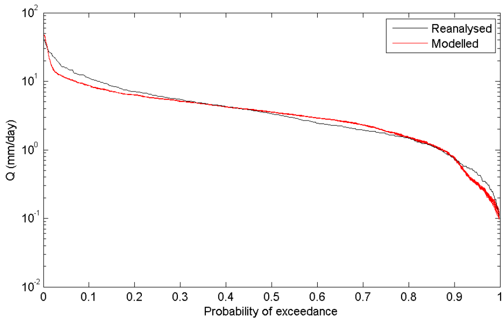


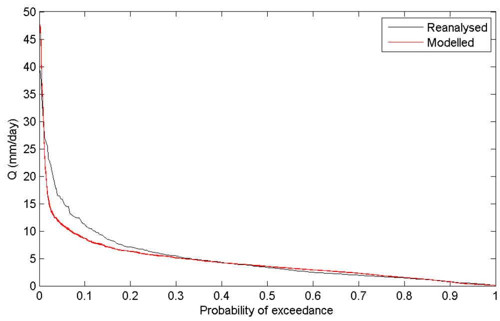


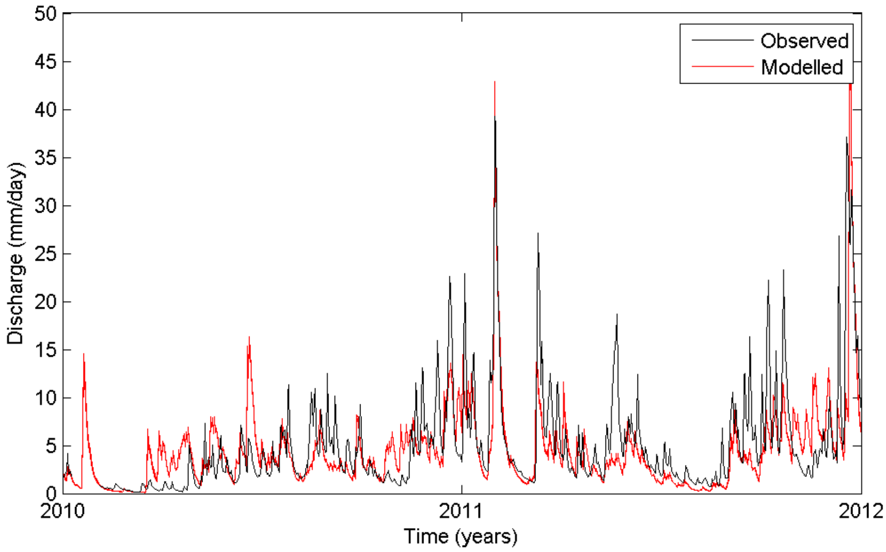
Sub-catchment 2, Topography driven, 3 parameters



Sub-catchment 2, Topography driven, 2 parameters





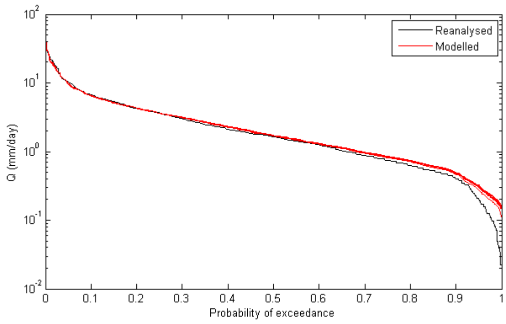


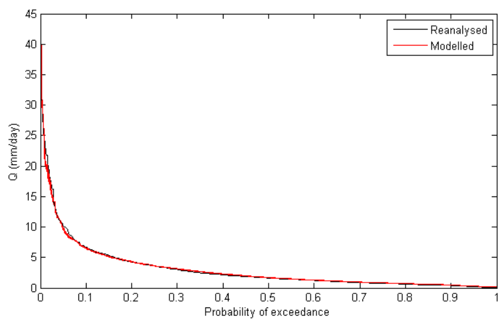
Appendix J – Hydrographs and flow duration curves during validation on a different catchment

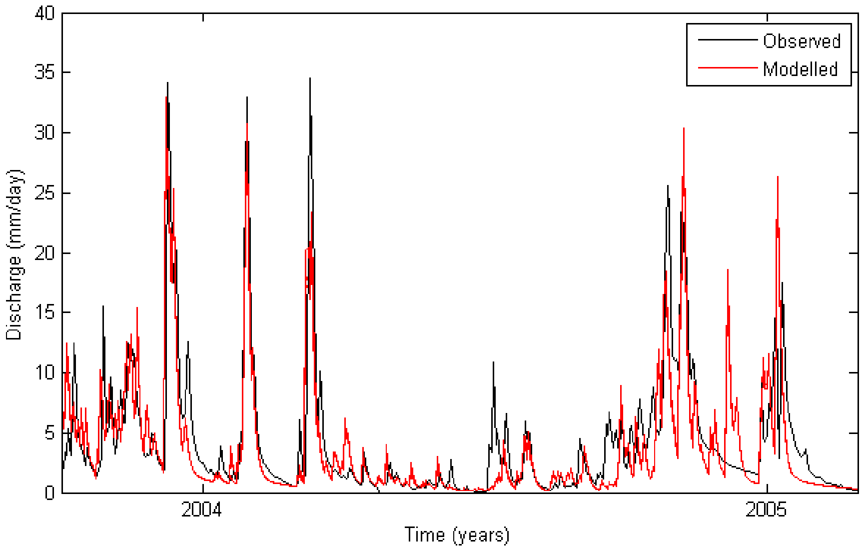
The figures below show the flow duration curves and hydrographs during validation on a different catchment. The modelling period was chosen similar to the period during calibration, being 2003-2007, whereby the first year was used as spin-up. The order of the figures is as follows: the first two figures describe the flow duration curve of the log of the flows and the flow duration curve of the flows. The third figure shows the hydrograph over a part of the validation period.

Topography driven
Sub-catchment 1 transferred to sub-catchment 2, 2 parameters

This model is relatively well capable of reproducing the hydrological signal of sub-catchment 2 (2-parameter definition). This model performed best compared to all models that were transferred to sub-catchment 2.

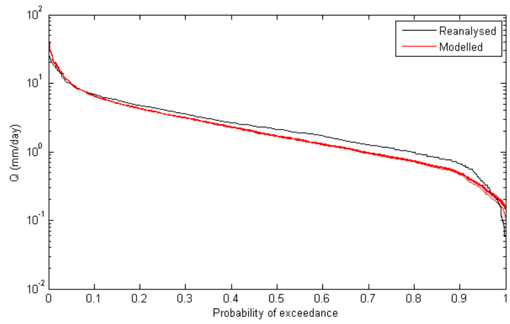


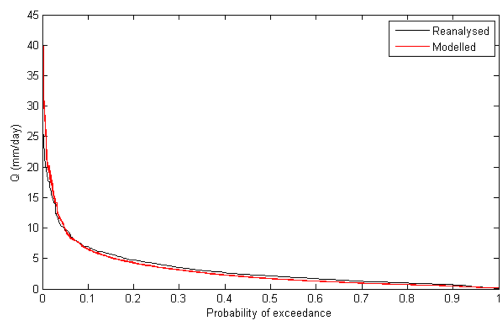


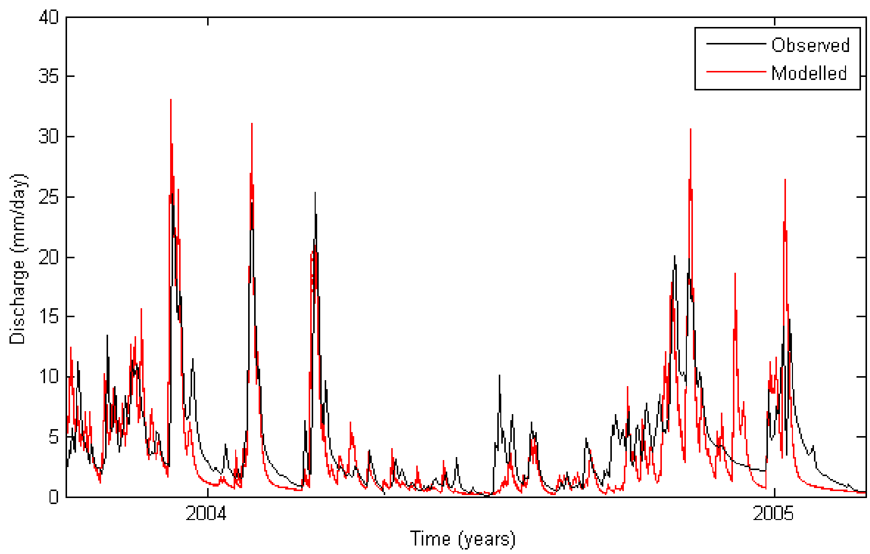


Topography driven
Sub-catchment 1 transferred to sub-catchment 2, 3 parameters

This model is relatively well capable of reproducing the hydrological signal of sub-catchment 2 (3-parameter definition), but overestimates most peaks and underestimates the low flows.

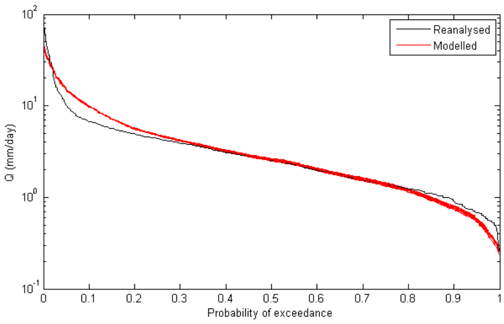


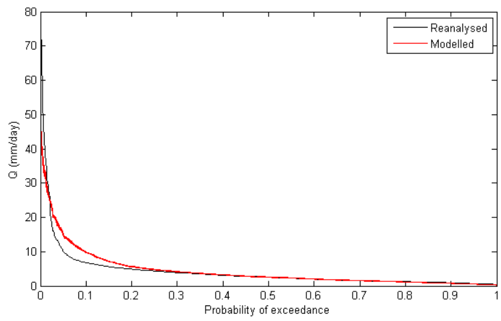


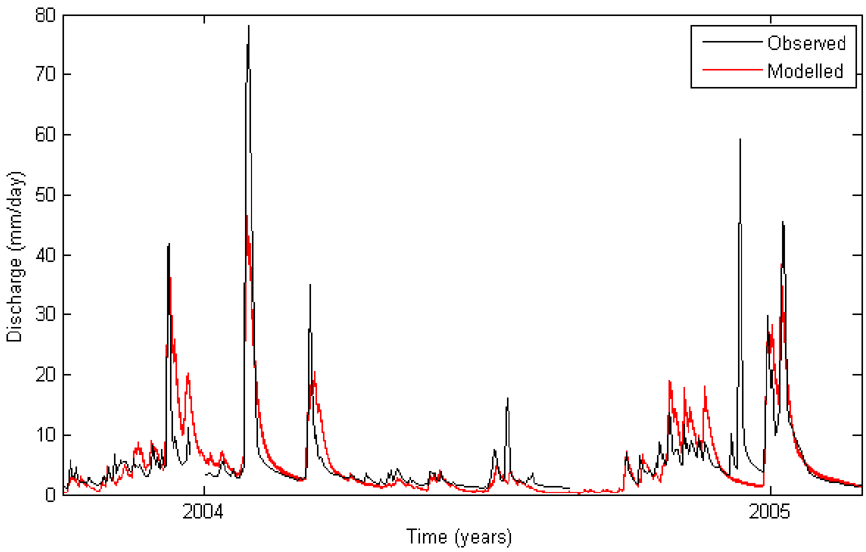


Topography driven
Sub-catchment 2, 2 parameters transferred to sub-catchment 1

This model shows relatively good performance in sub-catchment 1. It seems that 2 important peaks are missed, however, also in the calibrated model on this hydrograph these peaks were not simulated. This model performed best compared to all models that were transferred to sub-catchment 1.

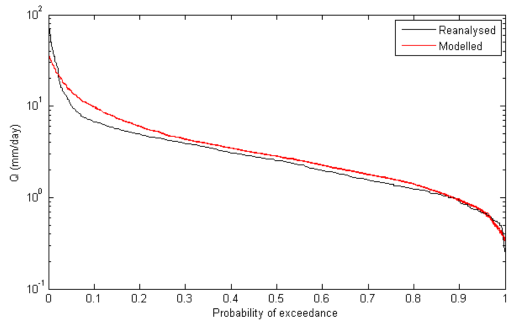


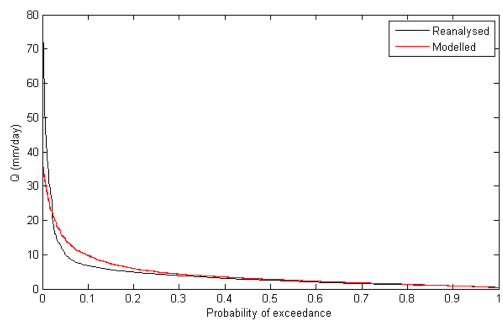


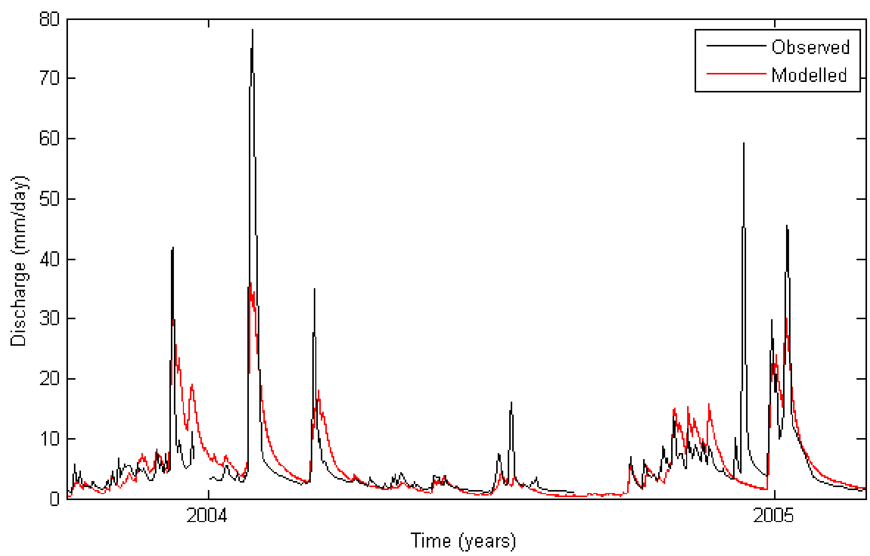


Topography driven
Sub-catchment 2, 3 parameters transferred to sub-catchment 1

This model shows relatively good performance in sub-catchment 1. It seems that 2 important peaks are missed, however, also in the calibrated model on this hydrograph these peaks were not simulated.

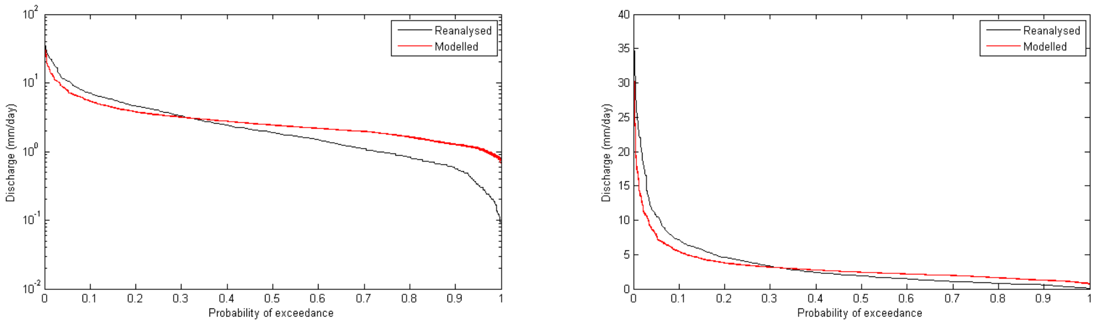


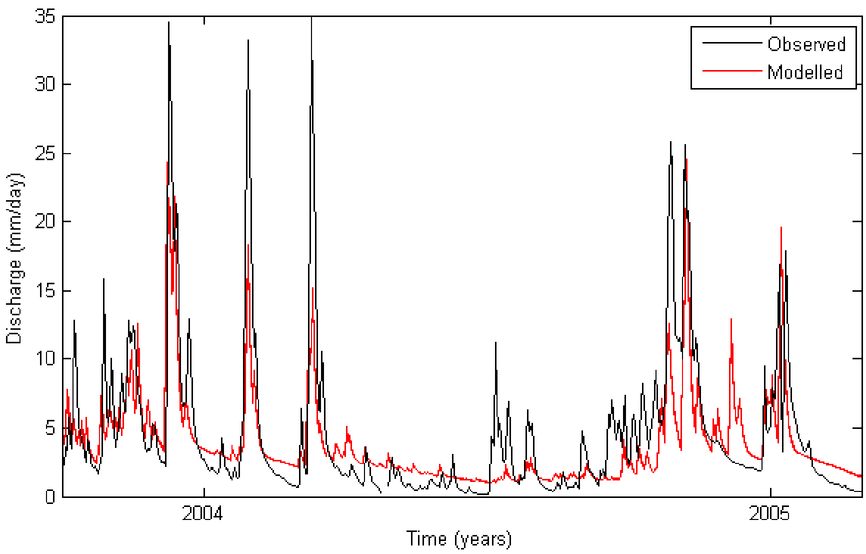




Lumped
Sub-catchment 1 transferred to sub-catchment 2, 2 parameters

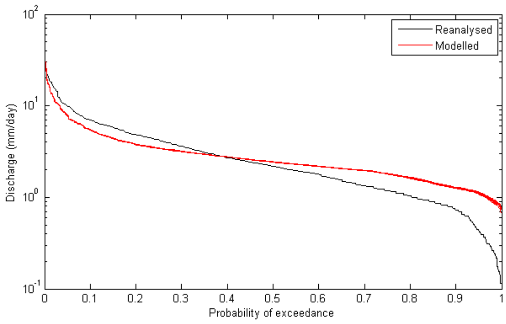
This model is not capable of reproducing the hydrological signal of sub-catchment 2 (2-parameter definition). The peaks are underestimated and the low flows overestimated.

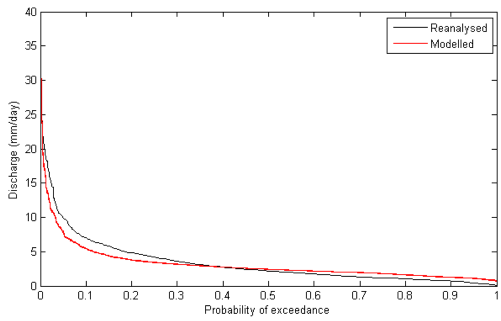


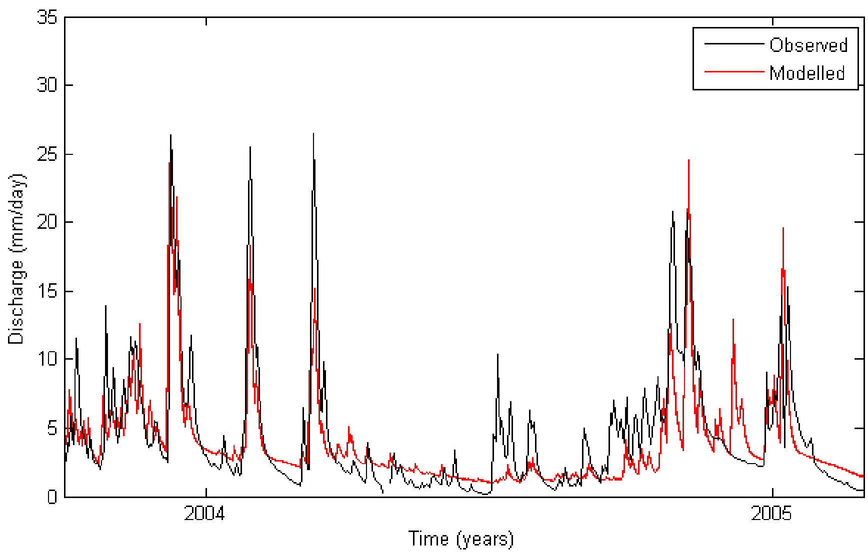


Lumped
Sub-catchment 1 transferred to sub-catchment 2, 3 parameters

This model is not capable of reproducing the hydrological signal of sub-catchment 2 (3-parameter definition). The low flows are overestimated.



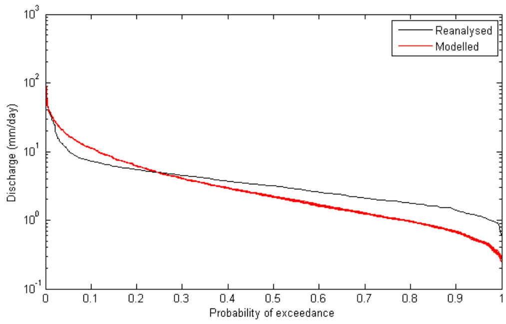


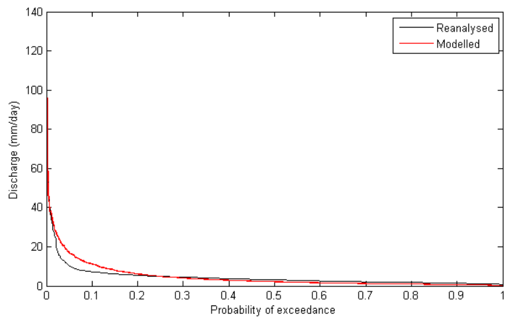


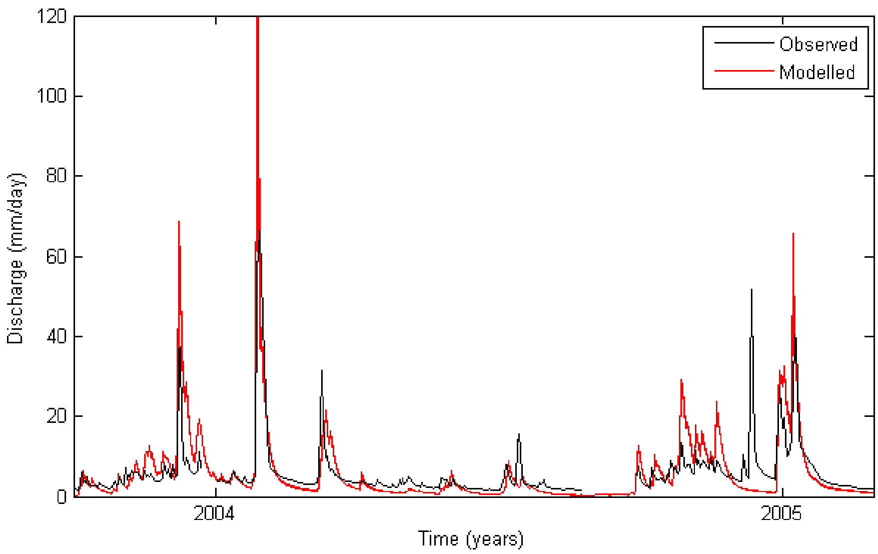
Lumped

Sub-catchment 2, 2 parameters transferred to sub-catchment 1

This model does not yield good results in catchment 1. The model overestimates the peaks to a large extend and underestimates the low flows.

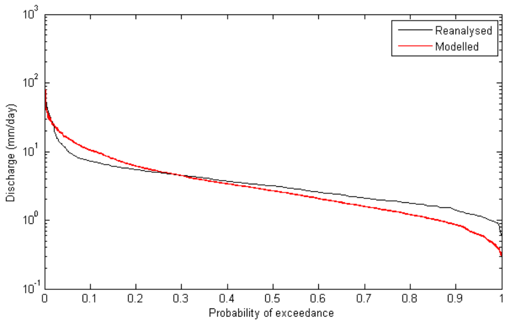


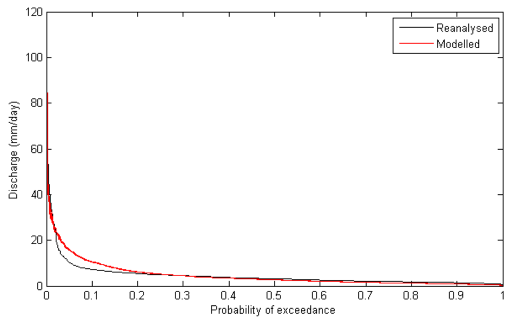


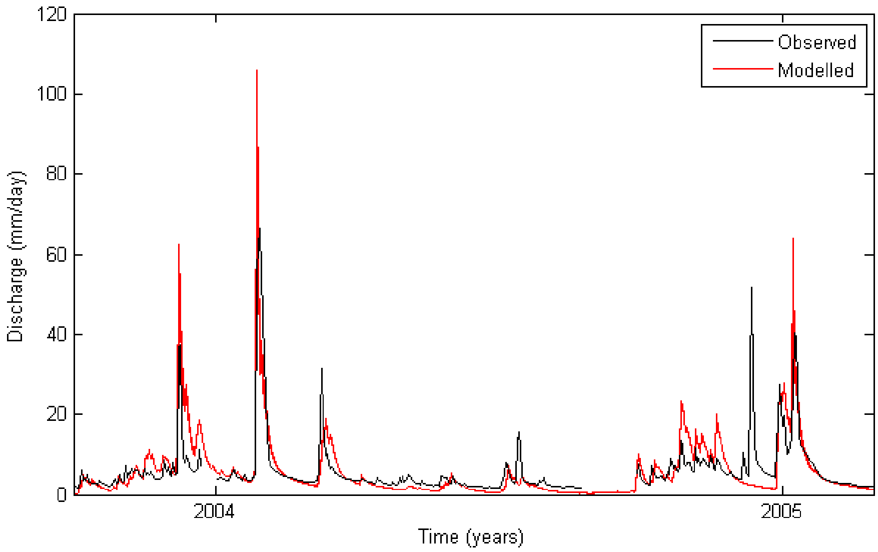


Lumped
Sub-catchment 2, 3 parameters transferred to sub-catchment 1

This model does not yield good results in catchment 1. The model overestimates the peaks to a large extend and underestimates the low flows.







Appendix K – Reanalysed long term annual averages

Table 10 shows the long term annual averages using the original measurement data, the reanalysed data and the data under the influence of a higher potential evaporation and a lower precipitation. Here, the long term annual averages were calculated over four years (2003-2007).
For the original data, the actual evaporation was determined as *P-Q*, as no proper models could be found to model the actual evaporation. For the reanalysed data, the actual evaporation has been determined using the modelled actual evaporation, which shows only slightly lower values than when *P-Q* was taken.

Table 10 - Original and reanalysed long term annual averages of the discharge, precipitation, runoff coefficient, actual and potential evaporation.

Catchment	type Q	Q	type P	P	RC	Eact	Epot
-	-	mm/year	-	mm/year	-	mm/year	mm/year
1	Original	1978	Gauge	2953	0,67	975	1443
	Original	1978	TRMM	2763	0,72	785	1443
	Reanalysed (lumped, 3 par)	1467	Gauge	2953	0,50	1365	1443
	Reanalysed (topo, 3 par)	1319	Gauge	2953	0,45	1423	1443
2	Original	743	Gauge	1932	0,38	1189	1438
	Original	743	TRMM	2820	0,26	2077	1438
	Reanalysed (lumped, 3 par)	1388	TRMM	2820	0,49	1303	1438
	Reanalysed (lumped, 2 par)	1398	TRMM	2820	0,50	1298	1438
	Reanalysed (topo, 3 par)	1349	TRMM	2820	0,48	1329	1438
	Reanalysed (topo, 2 par)	1303	TRMM	2820	0,46	1339	1438
2	Reanalysed (lumped, 2 par)	1020	TRMM	2820	0,36	1688	2157*
2	Reanalysed (lumped, 2 par)	299	TRMM	1410**	0,21	1038	1438

*Under the influence of 50% more potential evaporation
**Under the influence of 50% less precipitation

Appendix L – Steady state salt intrusion model equations

(54) $A = A_0 exp^{-\frac{x}{a}}$

(55) $B = B_0 exp^{-\frac{x}{b}}$

(56) $h = h_0 exp^{\frac{x(a-b)}{ab}}$

(57) $\frac{D}{D_0} = 1 - \beta \left(\exp\left(\frac{x}{a}\right) - 1\right)$

(58) $\frac{S-S_f}{S_0-S_f} = \left(\frac{D}{D_0}\right)^{\frac{1}{K}}$

(59) $L = a \ln\left(\frac{1}{\beta} + 1\right)$

(60) $D_0^{HWS} = 1400\frac{\bar{h}}{a}\sqrt{N_R}(\upsilon E)$

(61) $K = 0.2x10^{-3}\left(\frac{E}{H}\right)^{0.65}\left(\frac{E}{C^2}\right)^{0.39}(1-\delta b)^{-2.0}\left(\frac{b}{a}\right)^{0.85}\left(\frac{Ea}{A_0}\right)^{0.14}$

(62) $N_R = \frac{\Delta\rho}{\rho}\frac{gh}{\upsilon^2}\frac{Q_f T}{P_t}$

(63) $\beta = \frac{KaQ_f}{D_0 A_0}$

(64) $E = S\left(\frac{T}{2}\right) = \frac{\upsilon T}{\pi} = \frac{2\upsilon}{\omega}$

In which A (m^2) is the cross-sectional area, A_0 (m^2) is the cross-sectional area at the mouth, a (m) is the cross-sectional convergence length, B (m) is the stream width, B_0 (m) is the width at the estuary mouth, b (m) is the convergence length of the stream width, h (m) is the stream depth, h_0 (m) is the depth at the estuary mouth, x (m) is the distance towards upstream, D (m^2/s) is the longitudinal dispersion, D_0 (m^2/s) is the dispersion at the estuary mouth, β (-) is the dispersion reduction rate, K (-) is the dimensionless Van Den Burgh's Coefficient, Q_f (m^3/s) is the fresh water flushing, S (kg/m^3) is the steady state salinity, S_f (kg/m^3) is the fresh water salinity, S_0 (kg/m^3) is the salt water salinity, L (m) is the salt intrusion length, D_0^{HWS} (m^2/s) is the dispersion coefficient at high water slack, N_R (-) is the Richardson Number, υ (m/s) is the tidal velocity amplitude, H (m) is the tidal range, E (m) is the tidal excursion, C ($m^{0.5}/s$) is the Chézy coefficient, δ (-) is the damping number, $\bar{h}$ (m) is the tidal average stream depth, ω (s^{-1}) is the angular velocity, ρ (kg/m^3) is the density of the water, T (s) is the tidal period and P_t (m^3) is the flood volume (Savenije, 2012).

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., 1986. *An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system*. Journal of Hydrology, 87, 61–77.

Akaike, H., 1973. *Information theory and an extension of the maxium likelihood principle*. In 2nd International Symposium on Information Theory, Ed. B.N. Petrov and F. Csaki , pp. 267-81. Budapest: Akademia Kiado .

Aronica G.T., Candela A., Viola F. and Cannarozzo M., 2006. *Influence of rating curve uncertainty on daily rainfall-runoff model predictions*. In Predictions in Ungauged Basins: Promise and Progress, IAHS Publ. 303, Sivapalan M., Wagner T, Uhlenbrook S., Liang X., Lakshmi V., Kumar P., Zehe E., Tachikawa Y. (eds). 116–124.

Bergström, S., 1992. *The HBV model - its structure and applications*. SMHI Reports RH, No. 4, Norrköping.

Berrisford P., Dee, D.P., Fielding K., Fuentes, M., Kallberg, P., Kobayashi and S., Uppala, S.M., 2009. *The ERA-Interim Archive*. ERA Report Series, 1.

Berry, P.A.M., Garlich, J.D., Smith, R.G., 2007. *Near-global validation of the SRTM DEM using satellite radar altimetry*. Remote Sensing of Environment, Volume 106, Issue 1, Pages 17–27.

Beven K., 1993. *Prophesy, reality and uncertainty in distributed hydrological modelling*. Advances in Water Resources 16: 41–51.

Beven K., 1996. *Equifinality and uncertainty in geomorphological modelling*. In The Scientific Nature of Geomorphology, Rhoads BL, Thorn CE (eds). Wiley: Chichester, UK; 289–313.

Beven, K., 2000. *Uniqueness of place and process representations in hydrological modelling*. Hydrology and Earth System Sciences, 4, 203–213.

Beven K., 2001a. *How far can we go in distributed hydrological modelling?*. Hydrology and Earth System Sciences, 5(1): 1–12.

Beven, K., 2001b, *Rainfall-runoff modelling: The Primer*. John Wiley and Sons, Ltd.

Beven, K., 2002a. *Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modelling system*. Hydrological Processes, 16(2), 189–206.

Beven, K., 2002b. *Towards a coherent philosophy for environmental modelling*. Proc. Roy. Soc. Lond. A, 458, 2465–2484.

Blöschl, G. and Sivapalan, M., 1995. *Scale issues in hydrological modelling: a review*. Hydrological Processes, 9, 251-290.

Budyko, M.I., and Miller, D.H., 1974. *Climate and life*. volume 508, Academic press New York.

Cai, H., Savenije, H.H.G and Toffolon, M., 2012. *A new analytical framework for assessing the effect of sea-level rise and dredging on tidal damping in estuaries*. Journal of Geophysical Research, 117, C09023

Chan, N.W., 1997. *Increasing flood risk in Malaysia: causes and solutions*. Disaster Prevention and Management, Volume 6, Number 2, 72-86.

Chow, V.T., 1959. *Open-channel hydraulics*. International Student Edition, p111-113, McGraw-Hill Book Company, Inc.

Dee, D., Uppala, S., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A, Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Holm, E.V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M, Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F., 2011. *The ERA-Interim reanalysis: configuration and performance of the data assimilation system*, Quarterly Journal of the Royal Meteorological Society, 137, 553–597.

De Groen, M.M. and H.H.G. Savenije, 2006. *A monthly interception equation based on the statistical characteristics of daily rainfall*. Water Resources Research, 42, W12417.

Di Baldassarre G. and Montanari A., 2009. *Uncertainty in river discharge observations: a quantitative analysis*. Hydrology and Earth System Sciences Discussion 6: 39–61.

Euser, T., Winsemius, H.C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S. and Savenije, H.H.G., 2009. *A framework to assess the realism of model structures using hydrological signatures*. Hydrology and Earth System Sciences, 17, 1893–1912.

FAO, IUSS Working Group WRB, 2007, *World Reference Base for Soil Resources*. World Soil Resources Reports No. 103, Rome.

Fenicia, F., Savenije, H.H.G., Matgen, P. and Pfister, L., 2008. *Understanding catchment behavior through stepwise model concept improvement*. Water Resources Research, 44, 1.

Fischer, H.B., 1974. *Discussion of "Minimum length of salt intrusion in estuaries' by B.P. Rigter, 1973*. J. Hydraul. Div. Proc., ASCE, 100:708 712.

Gao, H., Hrachowitz, M., Fenicia, F., Gharari, S. and Savenije, H.H.G., 2013. *Testing the realism of a topography drives model (FLEX-Topo) in the nested catchments of the Upper Heihe, China*. Hydrology and Earth System Sciences Discussion, 10, 12663-1276.

Gharari, S., Hrachowitz, M., Fenicia, F. and Savenije, H.H.G., 2011. *Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment*. Hydrology and Earth System Sciences, 15, 3275–3291.

Gupta, H., Sorooshian, S. and Yapo. P.O., 1998. *Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information*. Water Resources Research, 34:751, 763.

Hrachowitz, M., Savenije, H.H.G., Bogaard, T.A., Tetzlaff, D. and Soulsby, C., 2013. *What can flux tracking teach us about water age distribution patterns and their temporal dynamics?*. Hydrology and Earth System Sciences, 17, 533–564.

ISRIC – World Soil Information, 2013. *SoilGrids: an automated system for global soil mapping.* Available for download at http://soilgrids1km.isric.org.

Kavetski D., Kuczera G., and Franks S.W., 2006a. *Bayesian analysis of input uncertainty in hydrological modelling: 1. Theory*. Water Resources Research 42(3): W03407.

Kavetski D., Kuczera G., and Franks S.W., 2006b. *Bayesian analysis of input uncertainty in hydrological modelling: 2. Application*. Water Resources Research 42(3): W03408.

Krueger T., 2009. *Uncertainties in modelling agricultural phosphorus transfers*. PhD thesis, Lancaster University.

Leonard J., Mietton M., Najib H. and Gourbesville P,. 2000. *Rating curve modelling with Manning's equation to manage instability and improve extrapolation*. Hydrological Sciences Journal - Journal Des Sciences Hydrologiques 45 (5): 739 – 750.

Liu Y.L., Freer J., Beven K. and Matgen P., 2009. *Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error*. Journal of Hydrology 367, 93–103.

Luxemburg, W.M.J. and Coenders, A.M.J., 2011. *Hydrological Measurements.* CIE4440 Hydrological Measurements, Delft University of Technology, unpublished.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T. and Clark, M., 2010. *Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions*. Hydrological Processes, 24, 1270–1284.

Mohamed, Y.A., Bastiaanssen, W.G.M., Savenije, H.H.G., van den Hurk, B.J.J.M. and Finlayson, C.M., 2012. *Wetland versus open water evaporation: An analysis and literature review*. Physics and Chemistry of the Earth, 47–48, 114–121.

Montanari A., 2004. *An attempt to quantify uncertainty in observed river flows: effect on parameterisation and performance evaluation of rainfall-runoff models*. AGU Fall Meeting Abstracts.

Monteith, J.L., 1965. *Evaporation and environment*. Symp Soc Exp Biol., 19, 205-34.

Moyeed, R.A. and Clark, R.T., 2005. *The use of Bayesian methods for fitting rating curves, with case studies*. Advance in Water Resources 28: 807–818.

NASA, 2012. *About TRMM*. Available online at [http://trmm.gsfc.nasa.gov, accessed on 01.09.2013].

Nash, J.E., and Sutcliffe, J.V, 1970. *River flow forecasting through conceptual models part i – a discussion of principles*. Journal of Hydrology, 10(3), 282-290.

Nguyen, A.D., and H.H.G. Savenije, 2006. *Salt intrusion in multi-channel estuaries: a case study in the Mekong Delta, Vietnam*. Hydrology and Earth System Sciences 10, 5.

Nijzink, R.C., 2013. *Improving the predictive equation for dispersion in estuaries*. MSc Thesis, Delft University of Technology.

O'Callagan, J.G. and Mark, D.M., 1984. *The extraction of drainage networks from digital elevation data*. Comput. Vision, Graphics, Image Process., 28, 323-344.

Pappenberger F., Matgen P., Beven K., Henry J.B., Pfister L. and Fraipont P., 2006. *Influence of uncertain boundary conditions and model structure on flood inundation predictions*. Advances in Water Resources 29: 1430–1449.

Penman, H.L., 1948. *Natural evaporation from open water, bare soil and grass*. Proceeding of the Royal Society of London, Series A, Mathematical and Physical Sciences, 193, 1032.

Petersen-Øverleir A., 2004. *Accounting for heteroscedasticity in rating curve estimates*. Journal of Hydrology 292(1–4): 173–181.

Pokhrel, P. and Gupta, H.V., 2011. *On the ability to infer spatial catchment variability using stream flow hydrographs*. Water Resources Research, 47.

Prandle, D., 1985. *On salinity regimes and the vertical structure of residual flows in narrow tidal estuaries*. Estuarine, Coastal and Shelf Science, 20, 615-635.

Refsgaard, J.C., Alley, W.M., and Vuglinsky, V.S., 1989. *Methodology for distinguishing between man's influence and climatic effects on the hydrological cycle*. IHP-I11 Project 6.3, Technical Documents in Hydrology, UNESCO, Paris.

Reitan T. and Petersen-Øverleir A., 2006. *Existence of the frequentistic estimate for power-law regression with a location parameter, with applications for making discharge rating curves*. Stochastic Environmental Research and Risk Assessment 20(6): 445–453.

Rennó, C.D., Nobre, A.D., Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J., and Waterloo, M.J., 2008. *HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia.* Remote Sens. Environ., 112, 3469–3481.

Sanmuganathan, K. and Abernethy, C.L., 1975. *A mathematical model to predict long term salinity intrusion in estuaries*. Proc. Second World Congress, International Water Resources Association, New Delhi, Vol. 3, pp. 313-324.

Savenije, H.H.G., 1986. *A one-dimensional model for salinity intrusion in alluvial estuaries*. Journal of Hydrology, 85 (1-2), 87-109.

Savenije, H.H.G., 1989. *Salt intrusion model for high-water slack, low-water slack and mean tide on spreadsheet*. Journal of Hydrology, 107, 9-18.

Savenije, H.H.G., 1997, *Determination of evaporation from a catchment water balance at a monthly time scale.* Hydrology and Earth System Sciences, 1, 000-000.

Savenije, H.H.G., 2001. *Equifinality, a blessing in disguise?.* Hydrological Processes, 15, 2835–2838.

Savenije, H.H.G., 2003. *Predictive model for salt intrusion in estuaries*. Journal of Hydrology, 148, 1–4, 203–218.

Savenije, H.H.G., 2004. *The importance of interception and why we should delete the term evapotranspiration from our vocabulary*. Hydrological Processes, 18(8), 1507– 1511.

Savenije, H.H.G., 2009a. *Hydrological Modelling*. CIE4431 Hydrological Modelling, Delft University of Technology, unpublished.

Savenije, H.H.G., 2009b. *HESS Opinions" The art of hydrology"*. Hydrology and Earth System Sciences, 13, 2.

Savenije, H.H.G., 2010. *Topography driven conceptual modelling (FLEX-Topo)*. Hydrology and Earth System Sciences, 14, 2681 – 2692.

Savenije, H.H.G., 2012. *Salinity and Tides in Alluvial Estuaries.* Second completely revised edition, http://hubertsavenije.files.wordpress.com/2014/01/salinityandtides2_21.pdf.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P.A. and Carrillo, G., 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. Hydrology and Earth System Sciences, 15, 2895–2911.

Schoups, G., Hopmans, J.W., Young, C.A., Vrugt, J.A., and Wallender, W.W., 2005. *Multi-criteria optimization of a regional spatially distributed subsurface water flow model*. Journal of Hydrology, 311, 20–48.

Schoups, G., Van de Giesen, N.C., and Savenije, H.H.G., 2008. *Model complexity control for hydrologic prediction*. Water resources research, 44, 12.

Shrestha, R.R., Bardossy A. and Nestmann F., 2007. *Analysis and propagation of uncertainties due to the stage-discharge relationship: a fuzzy set approach*. Hydrological Sciences Journal— Journal Des Sciences Hydrologiques 52 (4): 595 – 610 Shaw, E. M. 1994. *Hydrology in practice*, 3rd ed., Stanley Thornes (Publishers) Ltd, Cheltenham, UK.

Sivapalan, M., 2003. *Prediction in ungauged basins: A grand challenge for theoretical hydrology.* Hydrological Processes, 17, 3163-3170.

Smakhtin, V.U., 2001. *Low flow hydrology: A review.* Journal of Hydrology, 240, 147– 186.

Smith, A., 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Methuen & Co., Ltd, London.

Suhaila, J. and Jemain, A.A., 2007. *Fitting daily rainfall amount in Malaysia using the normal transform distribution*. Journal of Applied Sciences 7, 14, 1880-1886.

Tallaksen, L.M., 1995. *A review of base flow recession analysis*. Journal of Hydrology, 165, 349-370.

Thyer M., Renard B., Kavetski D., Kuczera G., Franks S.W. and Srikanthan S., 2009. *Critical evaluation of parameter consistency and predictive uncertainty in hydrological modelling: A case study using Bayesian total error analysis*. Water Resources Research, 45, W00B14.

Van der Burgh, P., 1972. *Ontwikkeling van een methode voor het voorspellen van zoutverdelingen in estuaria, kanalen en zeeën*. Rijkswaterstaat Rapport, pp. 10 72.

Van Rijn, L.C., 2011. *Analytical and numerical analysis of tides and salinities in estuaries; part i: tidal wave propagation in convergent estuaries*. Ocean Dynamics, 61, 1719-1741.

Venetis C., 1970. *A note on the estimation of the parameters in logarithmic stage–discharge relationships with estimation of their error*. Hydrological Sciences Bulletin, 15, 105–111.

Vogel, R.M., and Fennessey N.M., 1994. *Flow duration curves: I. A new interpretation and confidence intervals*. J. Water Resour. Plan. Manage., 120(4), 485– 504.

Vrugt, J.A. Braak ter, C.J.F. Clark, M.P., Hyman J.M., and Robinson, B.A., 2008. *Treatment of input uncertainty in hydrologic modelling: Doing hydrology backward with Markov chain Monte Carlo simulation*. Water Resources Research, 44, W00B09.

Wagener, T., 2007. *Can we model the hydrological impacts of environmental change?*. Hydrological Processes, 21, 3233–3236.

Wang-Erlandsson, L., van der Ent, R.J., Gordon, L.J., and Savenije, H.H.G., 2014. *Contrasting roles of interception and transpiration in the hydrological cycle. Part I: Simple Terrestrial Evaporation to Atmosphere Model*. Hydrology and Earth System Sciences Discussion, to be published.

Westerberg, I., Guerrero, J.L., Seibert, J., Beven, K., and Halldin, S., 2011. *Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras*. Hydrological Processes, 25, 603 – 613.

Winsemius, H.C., Schaefli, B., Montanari, A. and Savenije, H.H.G., 2009. *On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information*. Water Resources Research, 45.

Yilmaz, K.K., Gupta, H.V. and Wagener, T., 2008. *A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model*. Water Resources Research, 44, W09417.