MASTER THESIS

Predicting discharge of ungauged basins in the Brazilian Cerrados

Sebastiaan Agerbeek

supervised by

Drs. Roel Dijksma
Dr. Ir. Jos van Dam
Dr. Lineu Rodrigues

May 10th, 2016

Federal University of Viçosa

Wageningen University and Research Centre

Embrapa Cerrados

Predicting discharge of ungauged basins in the Brazilian Cerrados

Sebastiaan Agerbeek

May 10th, 2016

University	Wageningen University and Research Centre (Netherlands)
Master programme	Earth and Environment
Specialization	Hydrology and Quantitative Water Management
Course code	HWM-32806
First supervisor *Second supervisor*	Drs. Roel Dijksma Dr. Ir. Jos van Dam
External supervisors	Dr. Lineu Rodrigues, Embrapa Brazilian agricultural research corporation (Brazil) Prof. Fernando Pruski, University of Viçosa (Brazil)

Abstract

Rapid population growth in urban areas resulting in an increase in cropland and the use of hydro-electric dams are seen as causes of ecological problems in Brazil, of which the non-sustainable use of water resource is a major concern. Small rivers have an important ecological and socio-economic function, but are more vulnerable to environmental changes. This thesis tests a downscaling methodology, using a regional and a local watershed SWAT model, for the Preto river and the Buriti Vermelho river respectively. The models are constructed using data from the Brazilian Cerrados. This methodology attempts to predict discharge in other ungauged small rivers by using calibrated parameters from the regional watershed model as input for a local watershed model. The downscaling methodology resulted in a local watershed model able to predict average monthly discharge values with a Nash-Suthcliffe Efficiency > 0.1. This means that the discharge simulation using this downscaling methodology of ungauged river basins is only slightly better than the observed average discharge. Secondly, the regional watershed model was also used to assess the impacts of various environmental changes, namely climate change, land-use change and a hydropower dam on discharge. Climate change simulations showed a decrease in baseflow of the Preto river for the worst case scenario, with an annual discharge decrease of 18.1% in 2100. Land-use change mainly resulted in a decrease in baseflow and the hydropower dam caused a decrease in yearly discharge, as well as a decrease in seasonal fluctuations in discharge. These results show that these three different scenarios severely impact total discharge, but mostly groundwater flow in the Preto river watershed.

Keywords: *Catchment Hydrology, Soil Water Assessment Tool (SWAT), Predicting Ungauged Basins, Climate Change, Hydro-electric Dams, Land-use and Cover Change (LUCC), Brazil, Cerrados, Federal District, Preto river, Buriti Vermelho river*

Acknowledgements

I am very fortunate to be able to perform this research on hydrology and water resources in the Brazilian Cerrados and to write my thesis on the important topic of predicting ungauged basins and impact of (environmental) changes. This would not have been possible without the help of many people.

First of all, I want to thank the Wageningen University to give me the possibility to carry out the fieldwork and the writing on this fascinating topic. Most important for me were my supervisors, Roel Dijksma and Jos van Dam; their help and feedback helped greatly in organizing my data and writing down the results.

I was very lucky to travel to Brazil and am very grateful to have gotten this opportunity. In Brazil I was much welcomed by my local supervisor Lineu Rodrigues, whom I have to thank for introducing me to Brazil and the research, as well as for interesting conversations on both the field of hydrology and water resource management. Most of my time in Brazil I have spent at the Embrapa research institute, and I am thankful to all the people of Embrapa who I have met and who have been very supportive. Jose Roberto Rodrigues Peres has taught me the important research that is conducted by the Embrapa research institute. I would like to thank Luciano and Raquel too, for collecting and processing the data I needed for this research. Fernando Falco Pruski, from the University of Viçosa, has made it possible to study and live in Brazil for four months.

Lastly, I want to thank Maren for her support along the way. She has been of great help with data processing, GIS analysis, and much more. We had a great stay at the condominio and exploring Brasilía.

Table of contents

Chapter I

Introduction

1. Impacts on large and small rivers in the Federal District

Rapid population growth in urban areas results in an an increase in cropland for food and the use of hydro-electric dams for energy. These developments are seen as causes of ecological and hydrological problems in Brazil (Lorz et al., 2012). The capitol of Brazil is, since its construction in 1960 Brasília. This is an example of a large urban region where water demand, which is predicted to increase in the future, is reaching the limits of water supply. Brasilia lies in the Federal District (Distrito Federal = DF) of which the situation is concerning, because the population is expected to grow from currently 2.5 million to over 3.2 million in 2025. Figure 1 shows the population increase in the DF.

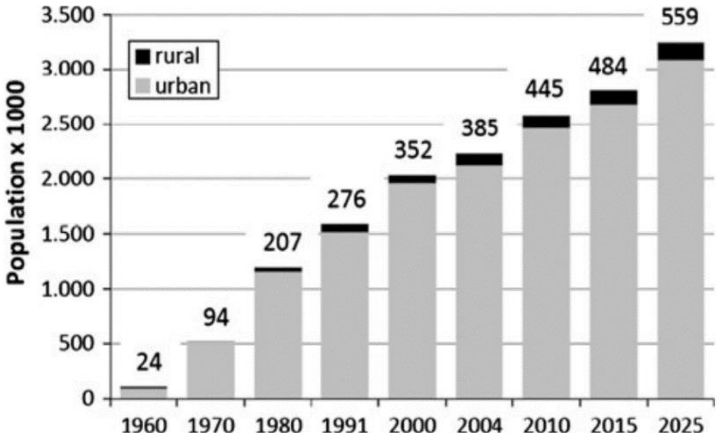


Figure 1. Historic and predicted population of the DF Brazil from 1960 to 2025 (Lorz et al., 2012)*; numbers above columns are population density (inhabitants per km²)*

Agriculture has increased from hardly any in 1954 to a 48% land coverage of the total DF in 2006. Figure 2 shows this increase, which mainly occurred from 1973 to 1998 and which resulted in more than 50% decrease of the savanna. Especially during the dry season, this large agricultural area is dependent on irrigation water, which is extracted from rivers surrounding the agricultural land (Lorz et al., 2012).

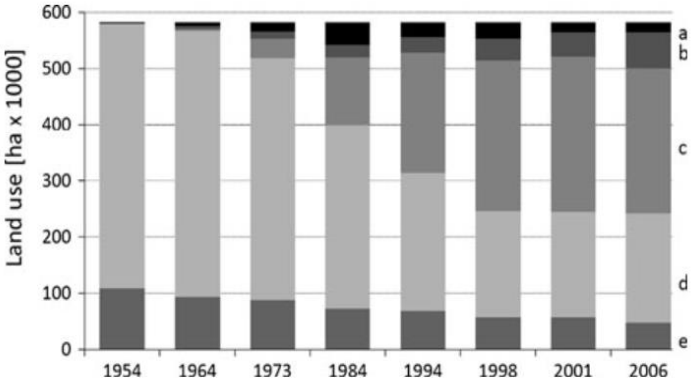


Figure 2. Land-use change from 1954 to 2006 (Lorz et al., 2012)*; **a** other land-uses, **b** settlements, **c** agriculture, **d** savanna, **e** forest*

Additional effects linked to the rapid population growth in the DF are a high water demand, resulting in surface water extraction from rivers used by small communities, urban areas and industrial areas. A high energy demand in the area incentivized building of hydro-electric dams, which have a large ecological impact on the downstream river watersheds (Soito & Freitas, 2011).

Lastly, the environmental change which is expected to impact water resources in the DF is climate change. For the DF, model predictions show decreased precipitation during the months June until September and increased temperatures of up to 4.5˚C , mainly during dry season from May until September (IPCC, 2008).

Surface water resources are provided by reservoirs, large rivers and small tributary rivers in the DF. An example of a large river is the Preto river, which delineates the eastern state boundary of the DF. The river is shown in figure 3A, close to Brasília. The Preto river has an important function for the state and the capital as provider of drinking water, irrigation water and is used to generate hydro-electricity. The Preto river itself is a large tributary river of the San Francisco river, of which the course through Brazil can be seen in blue in figure 3B.

In figure 3A, the light blue surface indicates the watershed of the Preto river, which is covered with thin blue lines which represent small tributary rivers. These small rivers have a large spatial coverage and are thus easily accessible and have an important function in providing water for farmers and small communities. Environmental change is expected to have a more severe impact on these smaller rivers, because they have a lower water capacity. Even though these small rivers have an important function and are vulnerable to environmental impacts, hydrological data is generally not collected from these sites. Assessment of environmental impacts on the water resources of these small rivers is very difficult. The Embrapa research institute has started collecting data from a small tributary river, the Buriti Vermelho river, indicated in figure 3C. The blue line in this figure 3C indicates the Buriti Vermelho river. The grey shapes indicate different land-use types, the dark grey circles are irrigated pivot agriculture lands. This proves this small river has an important function in providing water for irrigation.

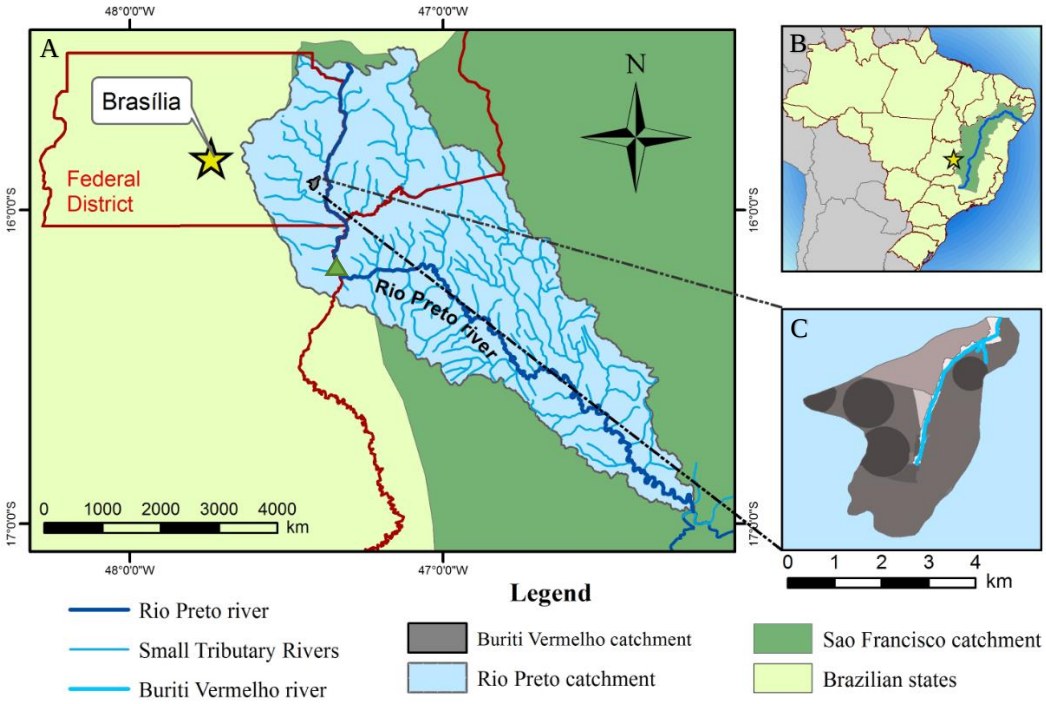


Figure 3. ***A*** *Figured on the left is a map of the Federal District, indicated is the location of Brasília, the FD state boundaries, the Preto river and catchment, the location of the Buriti Vermelho catchment (grey) and the location of a hydropower dam (green triangle).* ***B*** *In the upper right corner a world map showing Brazil in light green, the San Francisco basin (dark green) and indicating the location of Brasília (star).* ***C*** *The map in the right lower corner shows the Buriti Vermelho watershed and indicates the Buriti Vermelho river as a light blue line.*

The remainder of this chapter presents the research goals and questions, contains information about the study areas and describes the SWAT model. Chapter II describes the methodology, i.e. the method used for the prediction of ungauged small rivers and the scenario analysis, and the materials used. Moreover, in this chapter the collected data is described and how it is used to construct, calibrate and validate the SWAT model. Chapter III gives an overview and synthesis of the results. Finally, in chapter IV a conclusion and discussion is presented.

2. Research goals and questions

In this thesis, the problem of low hydrological data availability in small rivers is acknowledged. A methodology to simulate discharge of a local watershed using discharge data available from a regional watershed will be tested. Validation of this methodology requires the construction, calibration and validation of a regional watershed model, using the Soil Water Assessment Tool (SWAT model) (Arnold, Moriasi, et al., 2012). This model will additionally be used to assess various environmental changes in the regional watershed, e.g. land-use change, climate change and a hydropower dam. This results in two main goals:

1. Testing a methodology which can be used to simulate discharge in ungauged small rivers in the Cerrados, using the SWAT model;

2. Assessing the impact of various environmental changes, such as climate change, land-use change and hydro-electric dams on the hydrology of the Preto river, which borders the Federal District of Brazil.

In order to achieve the goals these research questions are formulated:

1. To what extent can discharge data from a large river be used to simulate discharge of a small river, using the SWAT model?
2. To what extent does land-use change impact discharge in the Preto river?
3. To what extent does climate change impact discharge in the Preto river?
4. To what extent does a hydro-electric impact discharge in the Preto river?

3. Study area and characteristics

Two study areas have been selected. The sub-watershed of the Preto river that borders the DF's eastern state boundary, which will be referred to as the *regional watershed*. The other study area is the Buriti Vermelho river watershed, which in this thesis is defined as the *local watershed*. The two study areas are located in the San Francisco basin (figure 3B), which is the third largest catchment in north-eastern Brazil. Only a small area of the San Francisco basin partly covers the DF). The total Preto river catchment is ±9700km2, the regional watershed model in this study only contains the upstream part of ±3300km2. The hydropower dam (Figure 3a) was built in the '90s. A large reservoir developed upstream of the dam, which has an area of 31 km^2. The Buriti Vermelho river is a tributary river of the Preto river (figure 3C). The area of this catchment is 9.2 km^2. The region contains a rural village with about 230 residents in 56 houses (van Vliet, 2012). Surrounding the river is a gallery forest. There are three agricultural areas that use pivot irrigation and there is an area destined for pasture. The key land-use classifications in the DF are presented in table 1. They are representative for both study areas, the Preto river and Buriti Vermelho river.

Table 1. Land-use types in the Federal District region defined by UNESCO (Lorz et al., 2012)

Class (engl./port.)	Description
Forest	
Mata	Forest of floodplains (*mata ciliar*)
	Gallery forest (*mata de galleria*)
	Dry forest (*mata mesofítica* (*seca*))
	Spare forest (*cerradão*)
Savanna	
Cerrado	Typical savanna (*cerrado típico*)
	Savanna with dense tree cover (*cerrado denso*)
	Savanna with sparse tree cover (*cerrado ralo/campo cerrado*)
Grass savanna and pastures	
Campo and Pastagem	Natural or managed grassland
Settlements	
Área urbana	Urban areas
Agriculture	
Área agrícola	Cropland, plantations, irrigated areas, service areas

De Souza Martins (2009) described the geology of the DF, which is representative for the two study areas. The DF lies on the Brasília Fold range, which consists of sedimentary rocks of the groups Paranoa (65%), Canastra (15%), Bambui (15%) and Araxa (5%). The topography and elevation differences in both study areas is determined by river erosion. The Buriti Vermelho river flows from south to north, where there is an elevation difference of 123 meters. The Preto river flows from north to south, with an elevation difference in the study area of 550 meters.

Both study areas have a tropical climate with average temperature of 21 °C and annual precipitation sums of between 1200 mm and 1750 mm. An important characteristic for the climate is the separation in dry and wet season. The wet season is from September until April in which 90% of the precipitation falls (De Souza martins, et al., 2009). Figure 4 shows monthly average precipitation and the reference evapotranspiration of the regional watershed. The graph shows that from approximately April until August the watershed experiences a water deficit, where reference evapotranspiration is higher than precipitation.

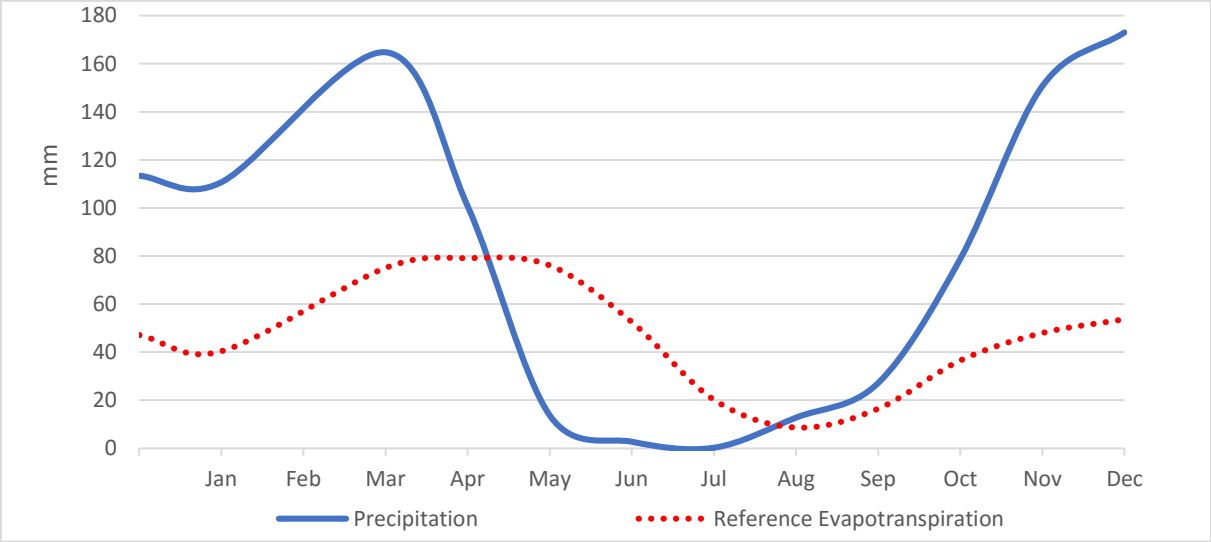


Figure 4. Monthly average precipitation and reference evapotranspiration (ET_{reff}) for the regional watershed. $ETreff$ is calculated using the Penman Monteith equation .

The soil of both watersheds is characteristic for the Cerrados, which is dominated by oxisols, cambisols and gleysols (Oliveira, Makeschin, *et al.*, 2014). Of the oxisols, the soiltype called latosol is the most dominant in both study areas. Latosols are red-colored soils of humid tropics. The formation is due to the process of leaching. The clay base is broken down by the warm humid environment. The remaining clay retains aluminum and iron compounds, resulting in the red color. Due to heavy rainfall and a high porosity nutrients in the soil are removed (Delvaux & Brahy, 2016). The soil depth is between 5 and 16 meters deep and had a high saturated hydraulic conductivity.

4. The SWAT model

The SWAT model is a water balance model based on physical processes, operating at a daily time step. It is designed to predict the impact of environmental changes on water, sediment, and agricultural chemical yields (Neitsch, Arnold, et al., 2009).

For this thesis, the SWAT model was chosen for several reasons. Terink, Lutz, *et al.* (2015) compared several hydrological models, of which the results can be viewed in table 2. The first two advantages named for the SWAT model are the most important. Firstly, in order to construct a model at a local and regional scale, it is important that SWAT allows different basin delineations at different scales. Secondly, this thesis presents an analysis of the impact of land-use change, for which the model's ability to incorporate changes in land-use and land cover are very relevant.

Table 2. Advantages and disadvantages of 6 different hydrological models when used for studying scale issues (Terink et al., 2015)

Model Name	Advantages	Disadvantages
SWB	+Easy to understand and implement +Requires almost no computational time +Data at lumped scale is available	-Small scale processes are omitted -No information on the spatial distribution of water in the basin
HBV	+The model structure is easy to understand +Requires little computational time +Model is freely available	-Requires manual basin delineation when dividing a larger basins in sub-basin -Interflows between basins are not accounted for
SWAT	+Allows for several different basin delineations at different scales +Can easily incorporate changes in land use and land cover +The model is open source and is freely available +Multiple calibration methods are readily available	-Requires more computational time, especially during calibration -Requires more data (suggesting that more assumptions will have to be made)
Pitman	+Has been successfully applied to semi-arid African basins before +The model structure is easy to understand	-Requires calibration of 24 parameters -The model is not well known
WEAP21	+Can incorporate the effects of water management decisions easily	-Does not focus on rainfall-runoff modelling -The model is licensed and will be difficult to obtain
GR4J	+The model structure is easy to understand +Requires calibration of only 4 parameters +Requires little computational time +The model code is readily available in MATLAB	-Requires manual basin delineation when dividing a larger basins in sub-basin

Another important justification for the choice to use SWAT is that over 100 studies have been performed in Brazil using SWAT (Bressiani *et al*., 2015). So, the SWAT model is an important decision support tool in Brazil. The developers have invested in giving workshops in Brazil, even all user documentation are available in Portuguese. The results from this thesis could thus be added to the large documentation studies using SWAT in Brazil.

In SWAT a catchment area is first divided into several subbasins, which are then divided further into hydrologic response units (HRUs). These consist of an area with homogeneous land-use, management, soil and slope characteristics. Precipitation landing on a HRU thus responds the same wherever it lands on the area of the HRU. The HRUs are not spatially identified within a SWAT simulation, because it is represented as a percentage of the subbasin area. The water balance of an HRU is subdivided into four storage reservoirs: snow, soil profile (0-2 meters), the shallow aquifer (2 – 15 meters), and the deep aquifer (more than 15 meters). The soil profile is also divided into the root zone, the unsaturated zone and the saturated zone. Within the soil profile various soil water processes are included, such as infiltration, soil evaporation, plant uptake, percolation to deeper layers and lateral flow. Flow from each HRU is then summed and then added to the channels or reservoirs until the watershed outlet. More detailed descriptions can be found in for example Arnold, Moriasi, et al., 2012; Inchell, Rinivasan, & Uzio, 2013; Neitsch et al., 2009.

Application of SWAT

Several steps can be identified to building a SWAT model. Figure 5 shows these steps as they were applied in building the two models for this thesis. The first step is to identify a watershed, in which a digital elevation map is needed in combination with a watershed boundary shape files, a shapefile showing the river network and a shapefile that indicates the subbasin distribution. The second step involves identifying the HRU's, for

which the SWAT model needs a land-use map, soil-map and slope map. Step 3 involves the SWAT program interpolating the available rainfall and weather data, for which SWAT needs data about the latitude, longitude and elevation of the hydrological and meteorological station and their time series of precipitation, humidity, solar radiation and wind speed. The SWAT program allows users to use a weather generator to generate weather data where data is missing. In step 4 the model can be run, calibrated using observed discharge data in step 5 and validated in step 6. If during the validation or calibration steps the model performs below a previously defined performance indicator value the model is rejected and previous steps must be redone. If the calibration and validation are accepted step 7 is reached and the model can be used for analysis.

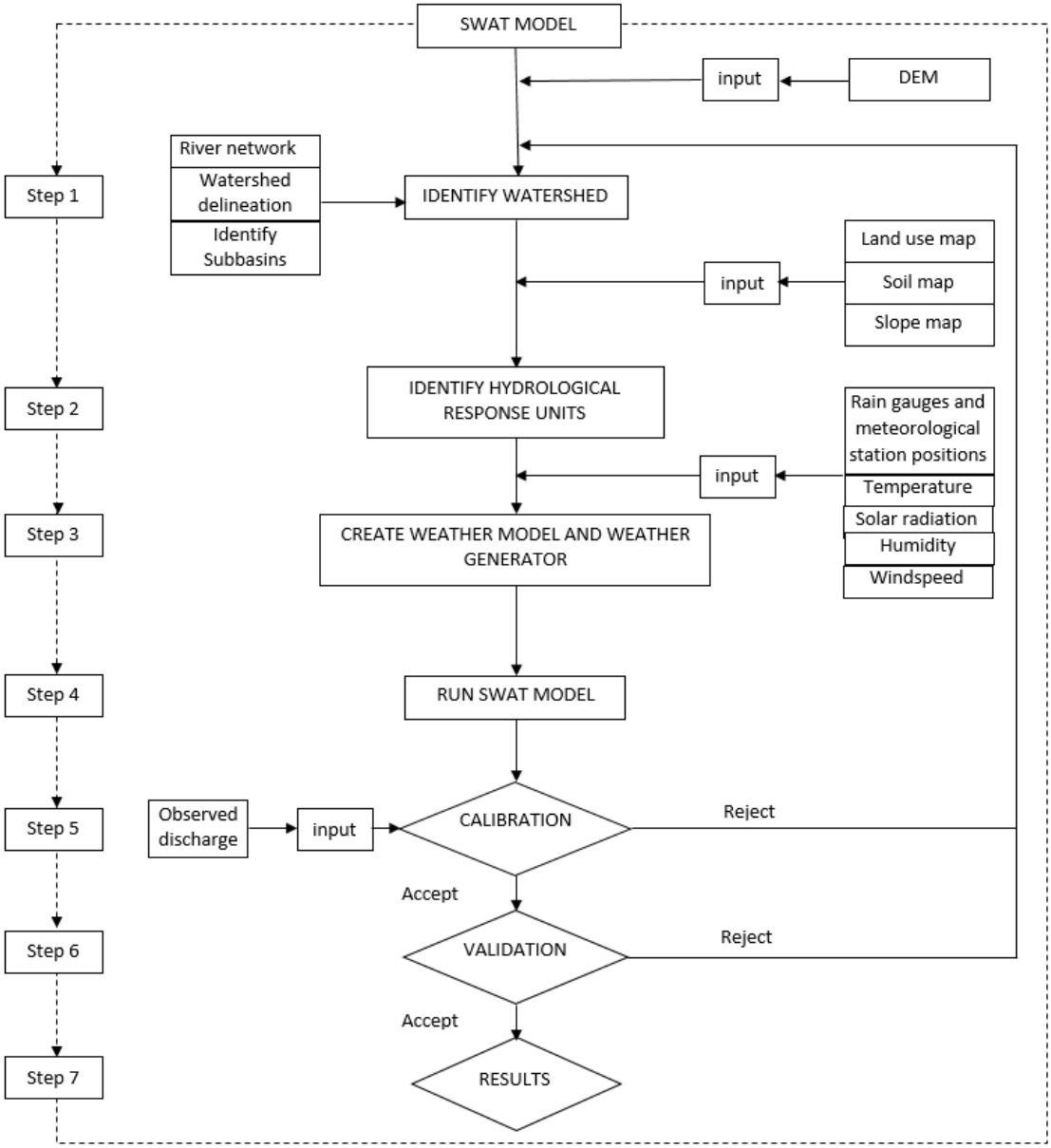


Figure 5. Conceptual framework of the SWAT application in this thesis

Chapter II

Materials and methodology

5. Prediction of discharge in an ungauged small river

In figure 6 it is conceptually explained how the regional watershed model was built, calibrated and validated in step 1, using geographical data prepared in ArcGIS, as well as weather data and discharge data from the Preto river as input data for SWAT. The discharge, weather and land-use data are from the period 1984 until 1994. During this period there was very little irrigated agriculture and the hydropower dam was not built yet. Therefore the discharge during this period is determined primarily by the natural hydrological system. The output of importance from the regional watershed SWAT model is a table of calibrated parameter adjustments. Step 2 was building a local watershed model, using geographical data and weather data from the period 2007 until 2014, because in this period discharge data was collected from the Buriti Vermelho river. In step 3 the calibrated parameter values from the regional watershed model are inserted in the local watershed model, which is referred to as the downscaling methodology. After step 3 a local watershed model with calibrated values originating from the regional watershed model is finished and can be tested. The goal is to simulate discharge of the Buriti Vermelho river for the period 2007 until 2015 and compare this to observed Buriti Vermelho discharge data. In this way the accuracy of the simulated discharge is assessed and the downscaling methodology is tested.

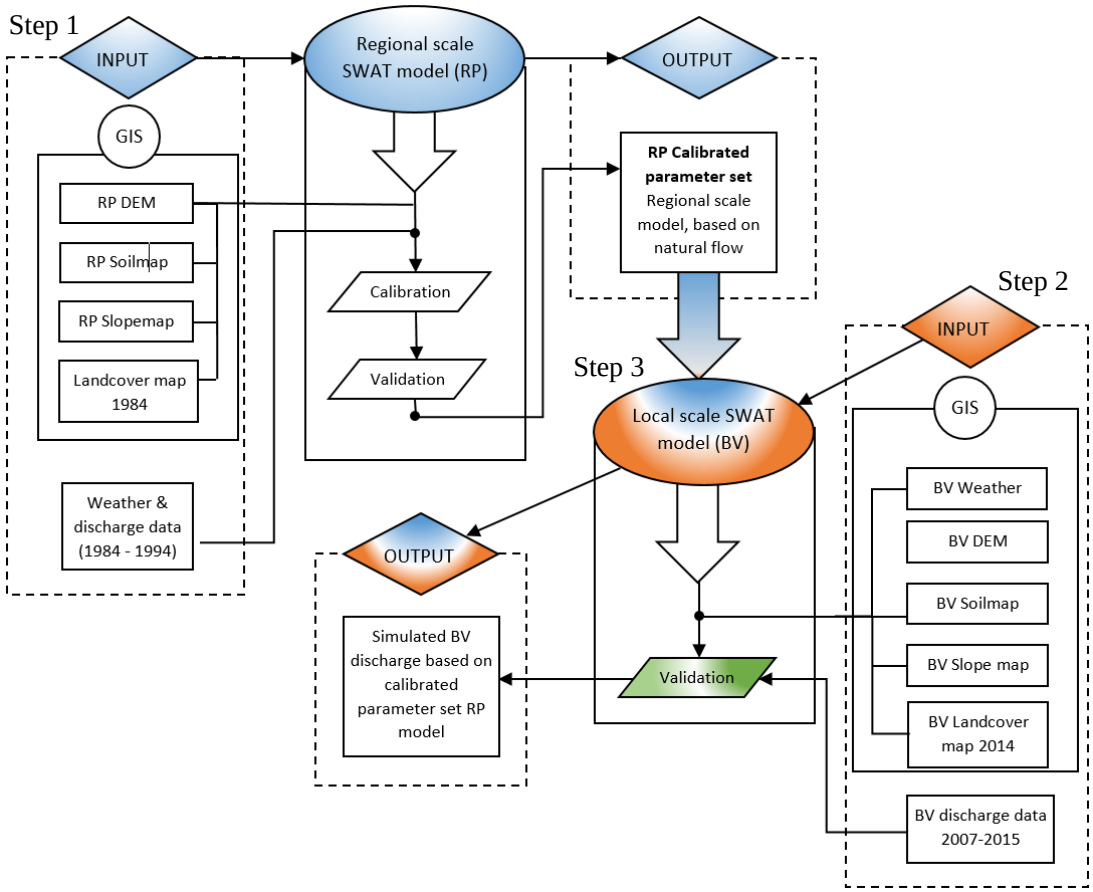


Figure 6. Conceptual description of the downscaling methodology, which attempts to predict discharge in local watersheds with the SWAT model using data from regional watersheds.

6. Scenario Analysis

6.1. Climate change impact assessment

The regional watershed model is used for climate scenario analysis. The climate scenarios were determined using input generated by the Coupled Model Intercomparison Project Phase 5 (Taylor, Stouffer, & Meehl, 2012), which is part of the IPCC Working Group on Coupled Modeling (IPCC, 2014). The CMIP5 studies the output of coupled atmosphere-ocean general circulation models and created the Global Climate Change Viewer (GCCV) (Alder, Hostetler, & Williams, 2013), which is an interactive web application that visualizes historical and simulated climate data. In figure 7, the simulated climate data for Brazil of temperature, is based on the CMIP5 input, which is based on the IPCC scenarios.

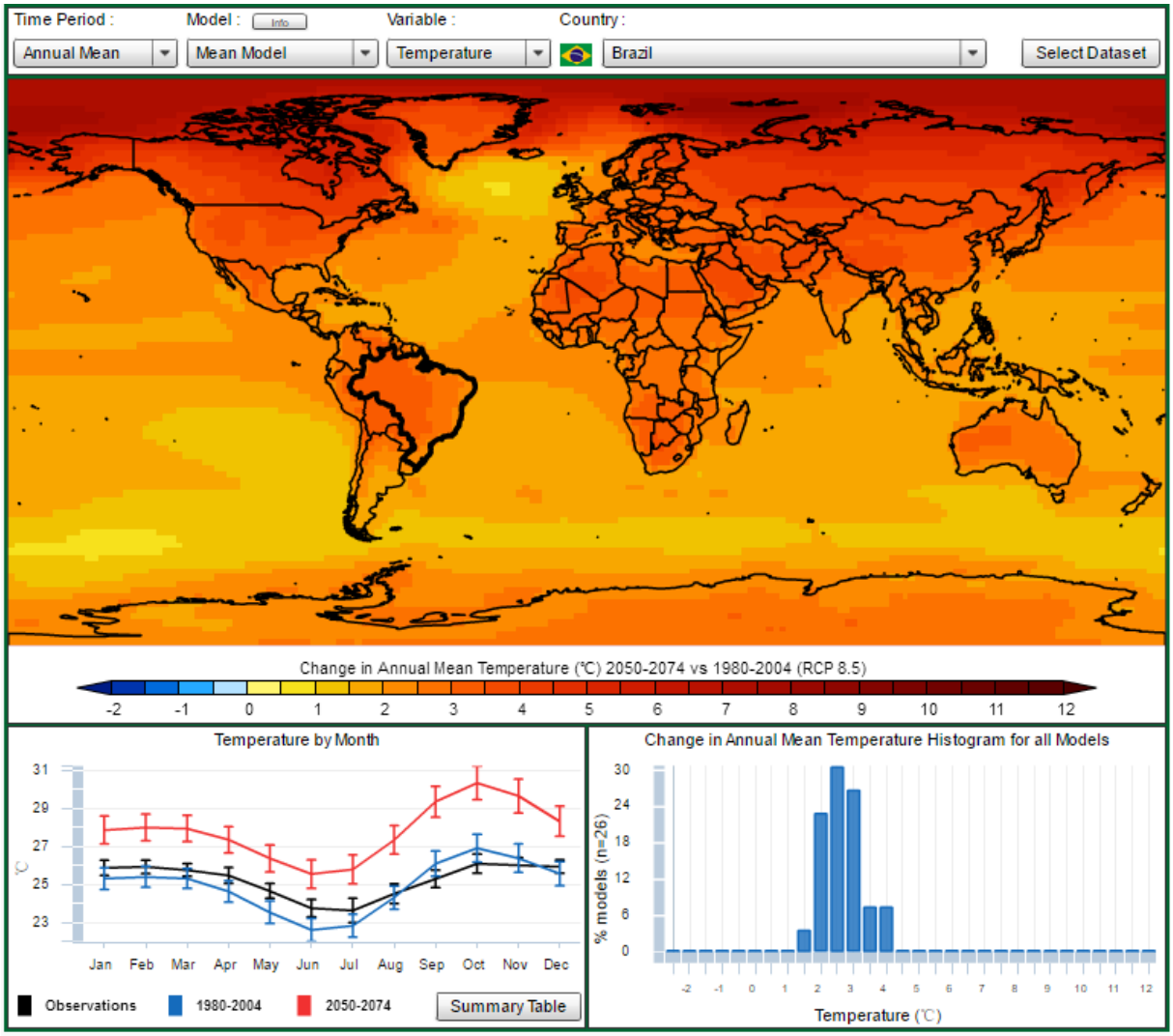


Figure 7. Screenshot of the Global Climate Change Viewer web application (Alder et al., 2013)

The SWAT model provides users with the ability to adjust the temperature and precipitation data used as input in the constructed SWAT models. This is done by inserting monthly percentage change in precipitation and temperature. Using the GCCV tool these percentage changes were collected to construct four climate scenarios, based on two IPCC Representative Concentration Pathways (RCP) emission scenarios; RCP4.5 and RCP8.5, for two years 2050 and 2100. These RCP scenarios focus on ambition levels of policy, which correspond to atmospheric greenhouse gas concentrations of 450, 650, 870 and 1400 ppmv CO_2-equivalent (IPCC, 2007). There are four emission scenarios, of which the names indicate the corresponding radiative forcing in the year 2100 in W^2/m^2 (IPCC, 2013). Figure 8 shows a conceptual model describing how climate change was assessed using the regional watershed model. After constructing, calibrating and validating the regional watershed model using geographical and hydro-meteorological data as input, the weather input data

was adjusted according to the four climate scenarios, resulting in discharge simulations for the four scenarios. By comparing the simulated discharge output time series to the observed discharge time series for the Preto river, climate change impact on discharge can be assessed.

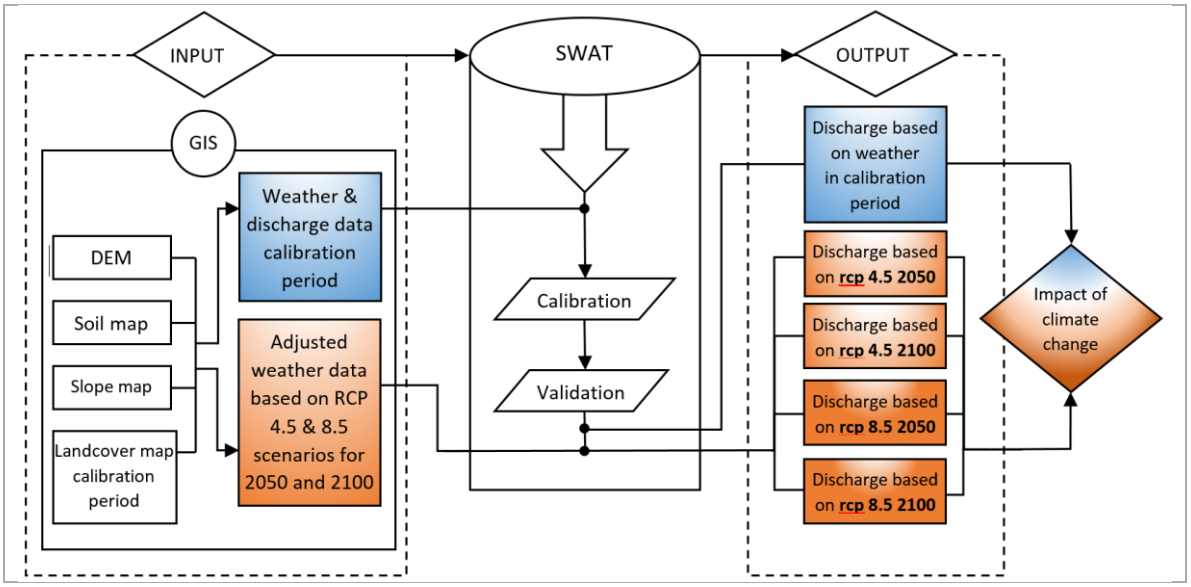


Figure 8. Conceptual description of the climate change scenario analysis performed with the regional watershed SWAT model for the Preto River.

Weather adjustments were made according to the two RCP scenarios for the two time periods in the regional watershed model. SWAT allows users to manually add percentage change for weather input data. It is possible to alter monthly precipitation, humidity, solar radiation, temperature data and carbon dioxide levels. SWAT input variables that can be altered for climate change analysis are shown in table 3.

Table 3. SWAT parameters used for climate scenario analysis. The table shows the variable name, definition and the input file in which the variable can be adjusted (Arnold, Kiniry, et al., 2012)

Variable Name	Definition	Input File
RFINC(mon)	adj_{pcp}: % change in rainfall for month	.sub
TMPINC(mon)	adj_{tmp}: increase or decrease in temperature for month (°C)	.sub
RADINC(mon)	adj_{rad}: increase or decrease in solar radiation reaching earth's surface for month (MJ m^{-2})	.sub
HUMINC(mon)	adj_{hmd}: increase or decrease in relative humidity for month	.sub
CO2	CO_2: carbon dioxide level in subbasin (ppmv)	.sub
IPET	Potential evapotranspiration method	.bsn

For the climate change assessment only precipitation and temperature data was altered, i.e. the input variables RFINC and TMPINC. Percentage increase or decrease has to be determined and entered per month, using the graphical interface shown in figure 9, which allows users to enter the percentage change values.

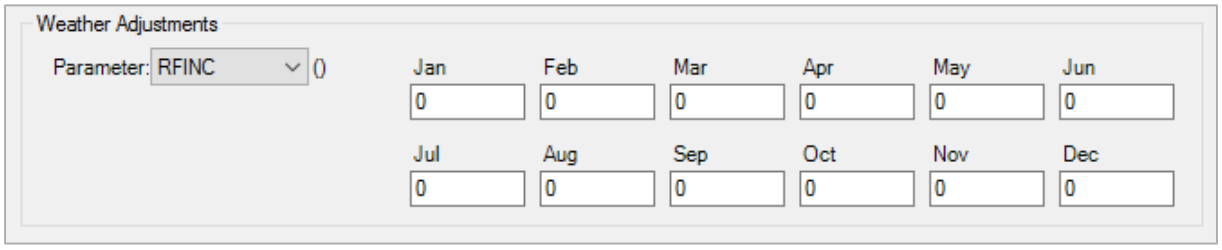


Figure 9. Screencap of the graphical user interface in arcSWAT, which allows users to adjust the parameters RFINC and TMPINC (Neitsch et al., 2009)

The SWAT model was run with the weather adjustment using weather and land-use data from the calibration and validation period, i.e. 1986 until 1994. Table 4 shows the monthly percentage changes made in the SWAT model for precipitation and temperature, which were determined using the GCCV tool. The

table is divided into two parts, one for precipitation and one for temperature. In both parts the middle row shows the historical average precipitation (mm/d) and temperature (°C). Above this row the simulated values are presented based on the four scenarios from the CMIP5 model, which takes the average of 20 global climate models (Alder et al., 2013). In the lower colored rows of the two sections (temperature and climate) the table shows the percentage change calculated as the difference between the historical monthly average and the simulated monthly change.

Table 4. Monthly precipitation and temperature values and their expected absolute and percentage changes compared to historical average monthly precipitation (mm/d) and temperature (Celcius). The colours indicate the severity of change. For precipitation the colour scale increases from green to yellow to red and for temperature the colour scale reaches from white to red. Values are based on the mean model of CMIP5.

Precipitation		**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
ΔP per month, period and RCP scenario(mm/d)	RCP 4.5 - 2050	0.2	0.1	0.1	0	0	0	0	-0.1	-0.1	-0.1	-0.2	0.1
	RCP 4.5 - 2100	0.2	0.1	-0.1	-0.1	0	0	0	-0.2	-0.2	-0.2	-0.2	0.1
	RCP 8.5 - 2050	0.2	0	-0.1	0	-0.1	-0.1	-0.1	-0.1	-0.3	-0.4	-0.3	0.1
	RCP 8.5 - 2100	0.1	0	-0.1	-0.1	-0.1	-0.1	-0.1	-0.2	-0.4	-0.5	-0.4	0
Historical average monthly precipitation in Brazil (mm/d)		6.8	6.8	6.7	5.4	3.4	1.9	1.1	1	1.6	2.8	4.6	6.2
Precentage difference per month, period and rcp scenario (%)	RCP 4.5 - 2050	2.94	1.47	1.49	0.00	0.00	0.00	0.00	-10.00	-6.25	-3.57	-4.35	1.61
	RCP 4.5 - 2100	2.94	1.47	-1.49	-1.85	0.00	0.00	0.00	-20.00	-12.50	-7.14	-4.35	1.61
	RCP 8.5 - 2050	2.94	0.00	-1.49	-1.85	-2.94	-5.26	-9.09	-10.00	-18.75	-14.29	-6.52	1.61
	RCP 8.5 - 2100	1.47	0.00	-1.49	-1.85	-2.94	-5.26	-9.09	-20.00	-25.00	-17.86	-8.70	0.00
Temperature		**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
T per month, period and RCP scenario(°C)	RCP 4.5 - 2050	27	27.2	27.1	26.4	25.3	24.4	24.6	26.2	28.2	29.1	28.5	27.4
	RCP 4.5 - 2100	27.2	27.4	27.3	26.7	25.5	24.6	24.8	26.4	28.4	29.4	28.8	27.6
	RCP 8.5 - 2050	27.9	28	27.9	27.3	26.4	25.5	25.8	27.3	29.3	30.3	29.6	28.3
	RCP 8.5 - 2100	29	29.1	29	28.5	27.6	26.8	27	28.6	30.7	31.8	30.9	29.5
Historical average monthly temperature in Brazil (°C)		25.3	25.4	25.3	24.6	23.5	22.6	22.8	24.3	26.1	26.9	26.4	25.6
Difference per month, period and rcp scenario (°C)	RCP 4.5 - 2050	1.7	1.8	1.8	1.8	1.8	1.8	1.8	1.9	2.1	2.2	2.1	1.8
	RCP 4.5 - 2100	1.9	2.0	2.0	2.1	2.0	2.0	2.0	2.1	2.3	2.5	2.4	2.0
	RCP 8.5 - 2050	2.6	2.6	2.6	2.7	2.9	2.9	3.0	3.0	3.2	3.4	3.2	2.7
	RCP 8.5 - 2100	3.7	3.7	3.7	3.9	4.1	4.2	4.2	4.3	4.6	4.9	4.5	3.9

The largest negative change in precipitation is predicted in the months August, September and October. Increases in precipitation are predicted in the months of December, January and February. These changes however are much smaller than the predicted decrease. The minimum temperature increases is 1.7%, decrease is not predicted in any months. The highest temperature increases are predicted in the months September, October and November. The maximum increase in 2100 predicted by the worst case RCP 8.5 scenario is 4.9% increase, i.e. 4.9 degrees warmer. Figure 10 shows the temperature and precipitation trends for the worst case scenario RCP 8.5 for the year 2100. The black line shows the historical average value and the

grey line the expected change. The red line has a secondary axis and indicates the percentage change. Both graphs show that change in precipitation and temperature is more prevalent in the months August, September and October, which is the second half of the dry period.

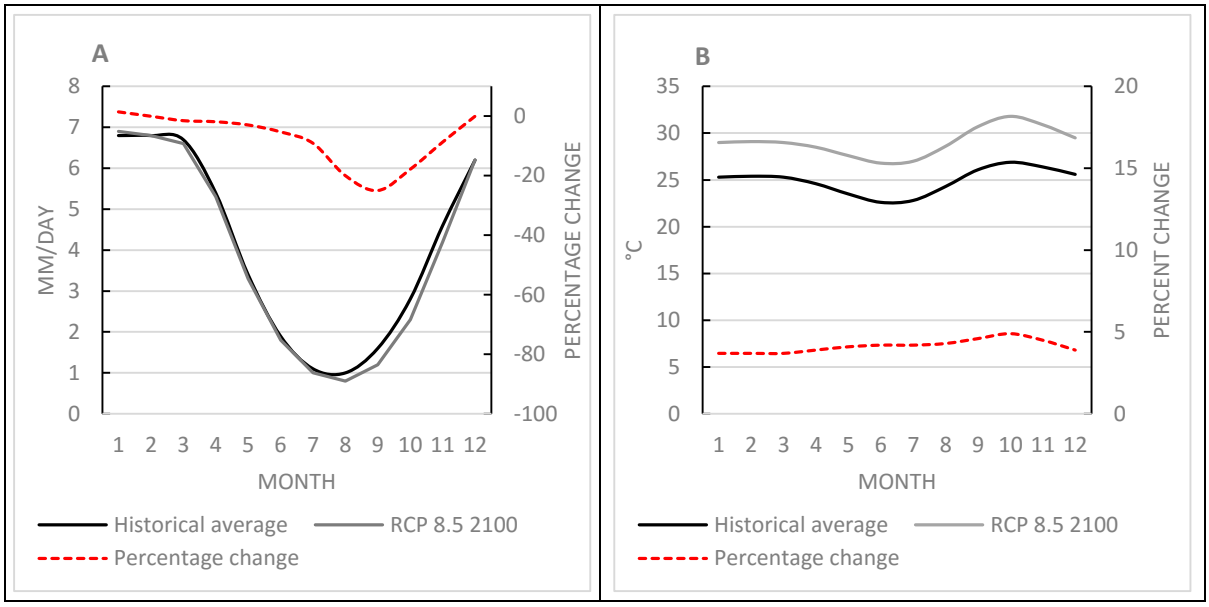


Figure 10. Monthly average precipitation (graph A) and temperature (graph B) values based on historical and RCP 8.5 IPCC scenarios in the Federal District. Figured values are determined using the Global Climate Change web application and correspond to the values in table 4. The historical and scenario monthly averages are respectively in black and grey. The red dotted line has a secondary axis, which represents the percentage change of the scenario values compared to the historical values.

6.2. Land-use change impact assessment

The Federal District has experienced a large shift in land-use and cover changes. Lorz et al. (2012) summarized how land-use has changed in the Federal District in Figure 11. The decrease in savannah *(d)* and the increase in agriculture *(c)* is the most prevalent observation.

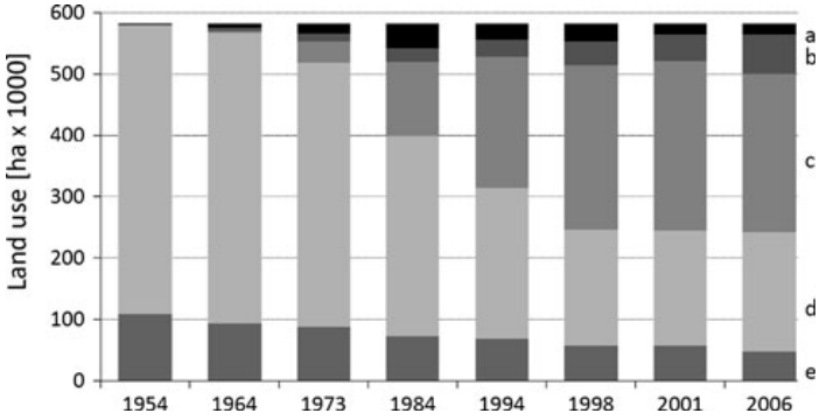


Figure 11. Land-use change from 1954 to 2006 (Lorz et al., 2012) ***a.*** *other land-use,* ***b****. settlements,* ***c****. agriculture,* ***d.*** *savanna, e. forest*

The regional watershed that is modelled has experienced a similar increase as was the case in the Federal District. The impact of land-use change on discharge can be analyzed in SWAT by changing the land-use map. The regional watershed model was calibrated using a land-use map from 1984. To assess how land-use change impacts the calibrated model the land-use map from 1984 is replaced by the land-use map from 2014. The precipitation and meteorological time series data is kept constant, to isolate the impact of land-use change. The results of total discharge, surface runoff and baseflow are analysed to draw conclusions. Table 5 shows how the main land-use types in the land-use map changed during the period between 1984 and 2014. A large part of the savannah, of ±850 km^2 has been replaced by agricultural land. Agriculture increased with

23.1 percent and savannah with 44.75%. The Queimados reservoir is also added to table 5 to show the extent of the area, which is bigger than the area used for urban settelments. There was no detailed agriculture map showing different crop types and agriculture types the choice was made to enter the 'agriculture land-generic' (AGRL) type option in arcSWAT. Figure 12 shows the land-use maps of 1984 and 2014 used in the regional watershed model.

Table 5. Land-use and cover change in the Preto watershed between 1984 and 2014 (Embrapa, 2016)

Year	1984	2014	1984	2014	
Land-use type	**area(km2)**	**area(km2)**	**Percent of total area(%)**	**Percent of total area(%)**	**Change (%)**
Agriculture	1433.7	2294.2	34.12	54.59	23.1
Urban	16.25	27.6	0.39	0.66	25.7
Savanna	2728.2	1880.5	64.92	44.75	-18.4
Queimados reservoir	0.0	31.8	0.0	0.3	∞
Pasture	24.1	0	0.57	0	∞

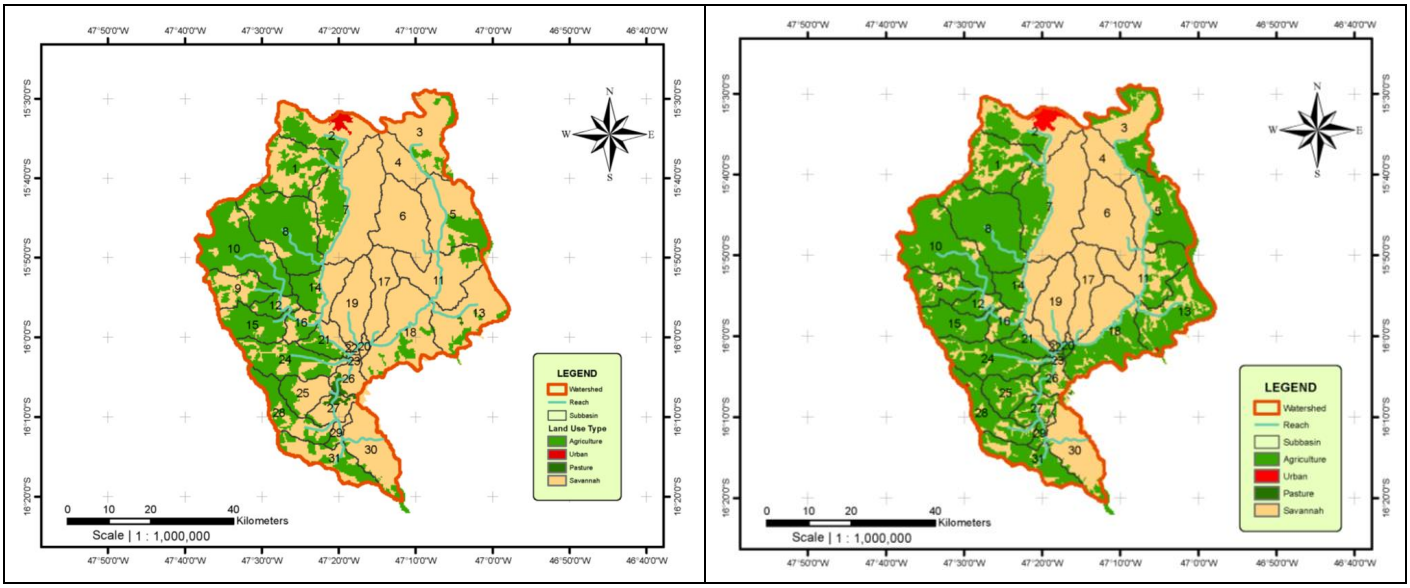


Figure 12. Land-use maps of regional watershed of 1984 (left) and 2014 (right) including identified subbasins. The red border indicates the watershed delineation. The polygons with the brown borders indicate the subbasins, which correspond to the numbers inside the polygons. Agriculture, urban settlements and Savannah are respectively indicated in green, red and beige.(Embrapa, 2016)

The land-use change methodology is conceptually explained in figure 13. After using the DEM, soil map, slope map and weather data as input for the SWAT model, calibration and validation is performed with the discharge data available during the calibration period. The regional scale model is calibrated during the period 1984-1994, because of the lack of anthropogenic influence on the hydrologic cycle. To assess the impact of change in land-use on the Preto river the land-use map of 1984, used during calibration, was exchanged for a land-use map of 2014.

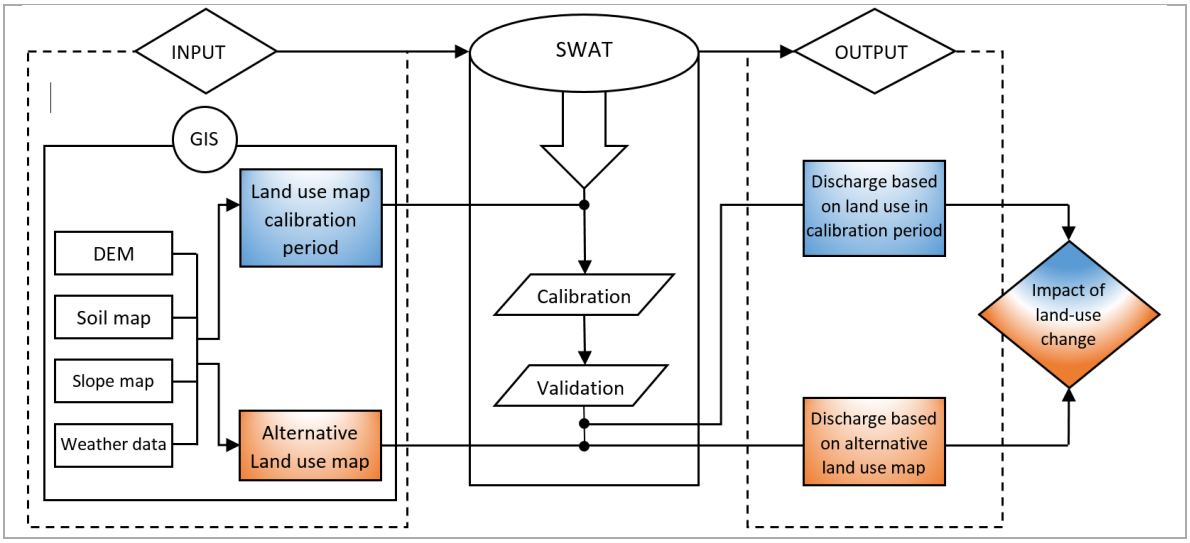


Figure 13. Conceptual description of the land-use change scenario analysis performed with the regional watershed SWAT model for the Preto River.

6.3. Hydro-electric dam impact assessment

Brazil is highly dependent on hydropower for its electricity, with over 80% of its electrical energy coming from large dams (Soito & Freitas, 2011). The hydropower dam placed in thePreto river caused the creation of the Queimado reservoir shown in figure 14.

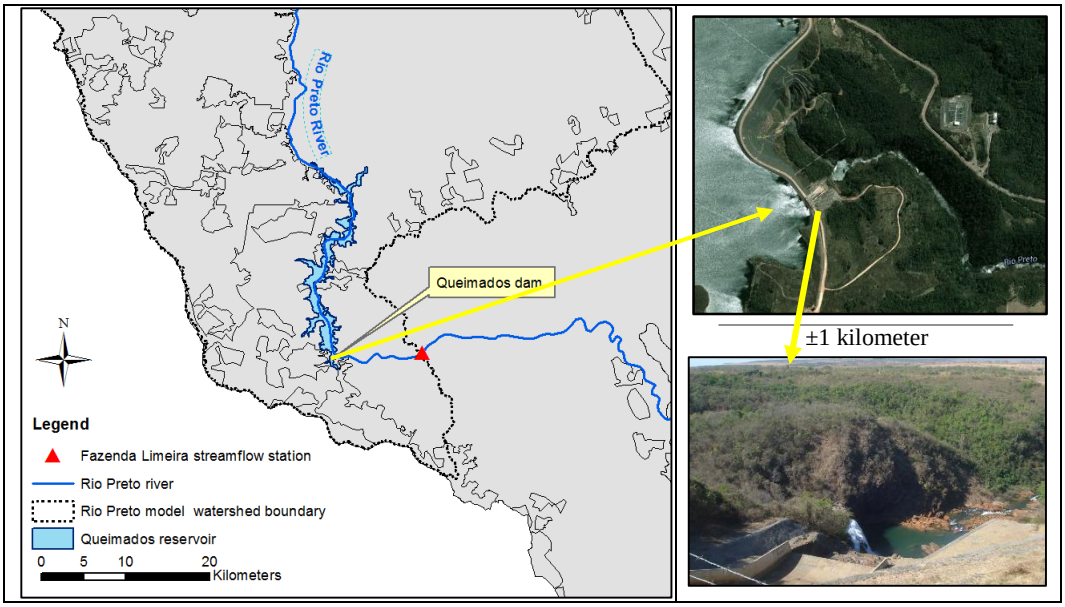


Figure 14. Map (left) of the Queimados reservoir, indicating the Preto river in blue, a hydrological measurement station located in Fazenda Limeira (red triangle) and the text box indicates the location of the dam. The image in the upper right is a screenshots of the dam from Google Earth (and the image in the lower right is taken by Luiz Paulo Oliveira (2011)

The impact of the hydropower dam on discharge was assessed for the regional watershed. The method is conceptually described in figure 15. The SWAT model was first built using geographical input data. The land-use map from 1984 is used. The model was then calibrated using discharge and weather data during the period 1984 and 1994, before the building of the Queimado hydroelectric power plant started in august 2000, which generates 105 MW (Cowley, 2003). After the calibration and validation of the regional model is completed the land-use data, weather and precipitation data is exchanged for land-use, weather and precipitation data from 2004 until 2014 allowing a simulation of discharge based on weather and land-use in this period, but without the influence of the Queimado dam. Comparing this simulated discharge with

observed discharge data from the Preto river downstream of the Queimados dam allows for the assessment of the impact of a hydroelectric dam on discharge.

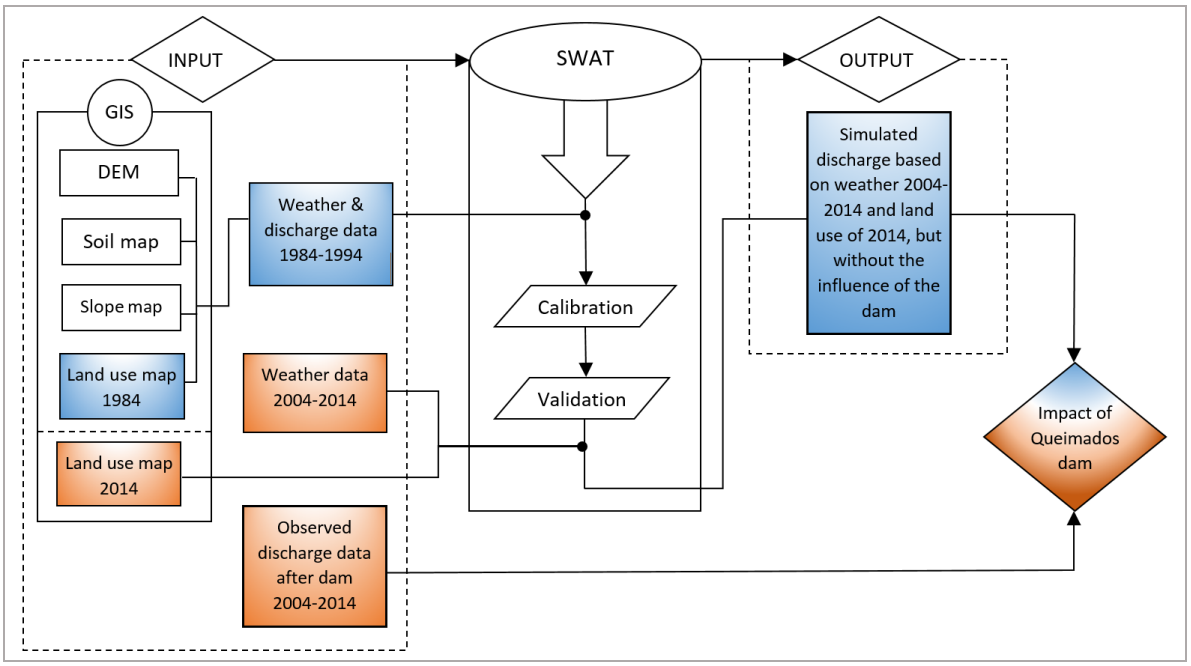


Figure 15. Conceptual description of the hydro-electric dam scenario analysis performed with the regional watershed SWAT model for the Preto River.

7. Materials

This chapter presents all the materials, in the form of time-series and geo-information maps that were used to build the regional watershed and local watershed SWAT models. The chapter is divided in two paragraphs, one for each model. Input data for both SWAT models was collected from local and global sources including a digital elevation map (DEM), land-use map, soil map, meteorological and hydrological data. Detailed information about the data can be found in appendix A.

7.1 Regional watershed model data

Geo-information data

DEM data was collected from the SRTM data set by NASA, which has a 30 meter resolution (Tang, Gao, Lu, & Lettenmaier, 2009) Figure 16 shows the elevation map, where the elevation is separated in six layers. The maximum elevation difference is 562 meters. Subbasin 30 contains the lowest elevation and east of this subbasin the elevation depletes even more. This elevation difference incentivized building the hydropower dam in subbasin 27.

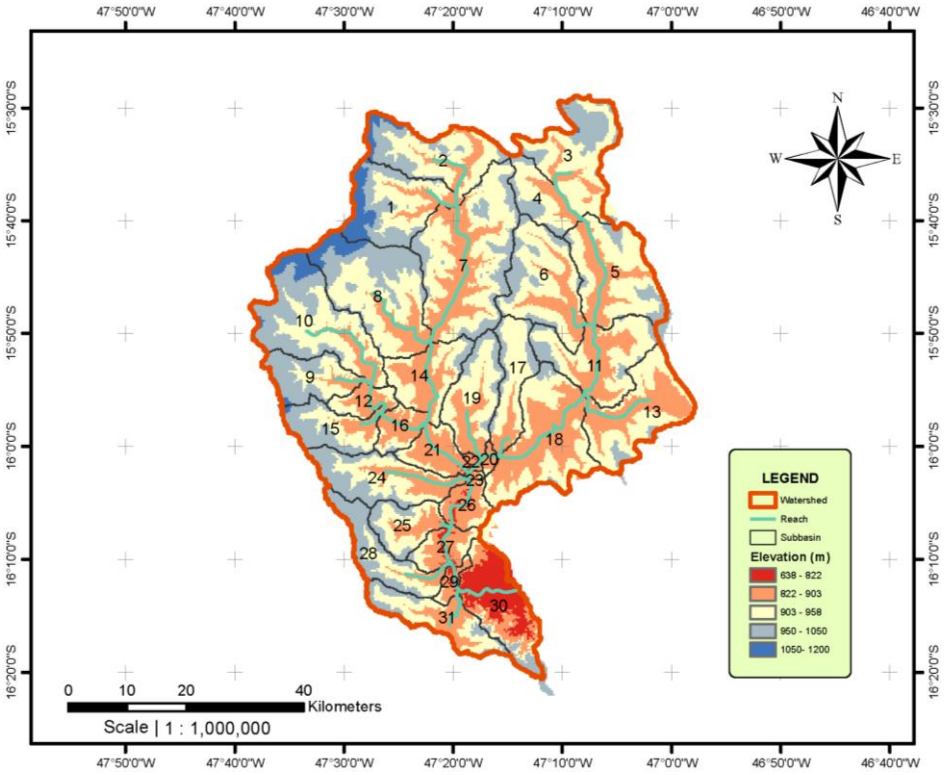


Figure 16. Digital Elevation Map of the regional watershed model, containing the Preto river. The red border around the area delineates the watershed boundary, there are five elevation levels, subbasins are indicated and the reach. (NASA, 2016)

The DEM data used in this research was chosen after comparison to other sources of DEM datasets, such as the DEM from the Instituto Brasileiro de Geografia e Estatistica (IBGE), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and the Shuttle Radar Topography Mission (SRTM). Gouvêa *et al*. (2005) compared the IBGE DEM to the SRTM DEM in a study of the Camanducaia basin in the State of São Paulo and concluded that the SRTM data showed more details. The SRTM DEM provides a higher horizontal resolution of 3 arc seconds compared to ASTER (Nikolakopoulos, *et al.* 2006). This justified the choice for the SRTM DEM from NASA.

The soil map was provided by the Embrapa research institute (Embrapa, 2016). The main soil types as can be seen is figure 17 and are variations on latosols and cambisol. The figure shows twelve soil types and indicates the subbasins of the area. The upper right region is mainly latosol and the upper left region is red latosol. These soil parameters used necessary for SWAT do not differ much from each other.

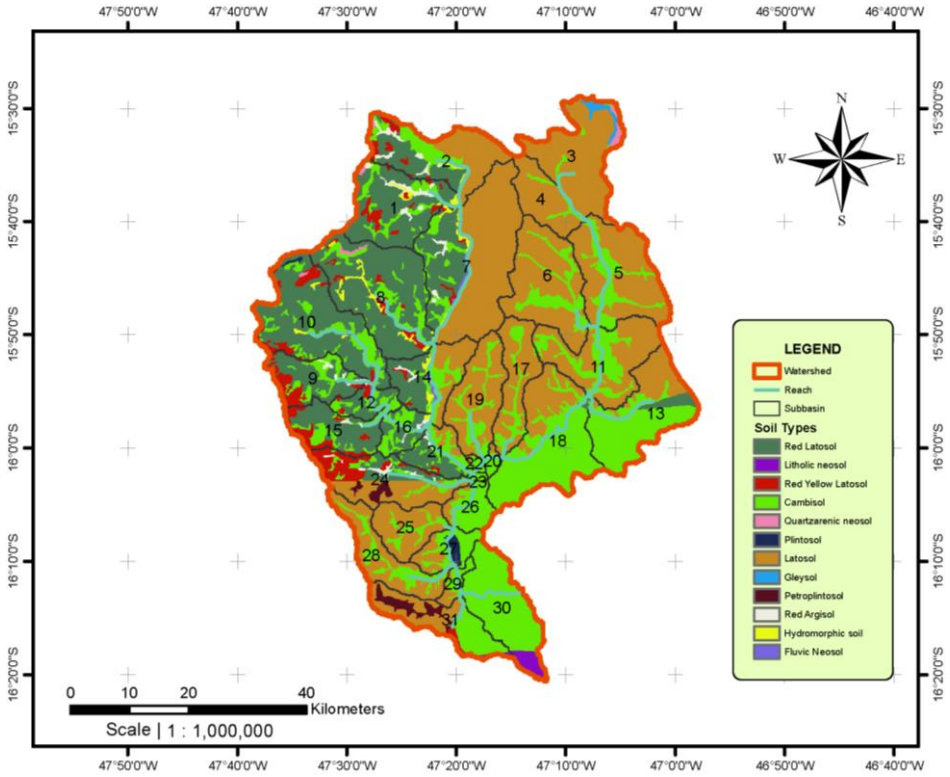


Figure 17. Soil map of the regional watershed model, containing the Preto river. The red border around the area indicates the watershed boundary, there are twelve soil types, sub-basis are indicated as wel as the reach. (Embrapa, 2016)

Land-use data was provided by the Embrapa research institute based on data collected from Landsat imagery (Embrapa, 2016). The land-use map used for the calibration of the regional watershed model is from 1986, shown in figure 18. Savanna is the main land-use type, covering more than 70% of the total area, followed by agriculture and pasture.

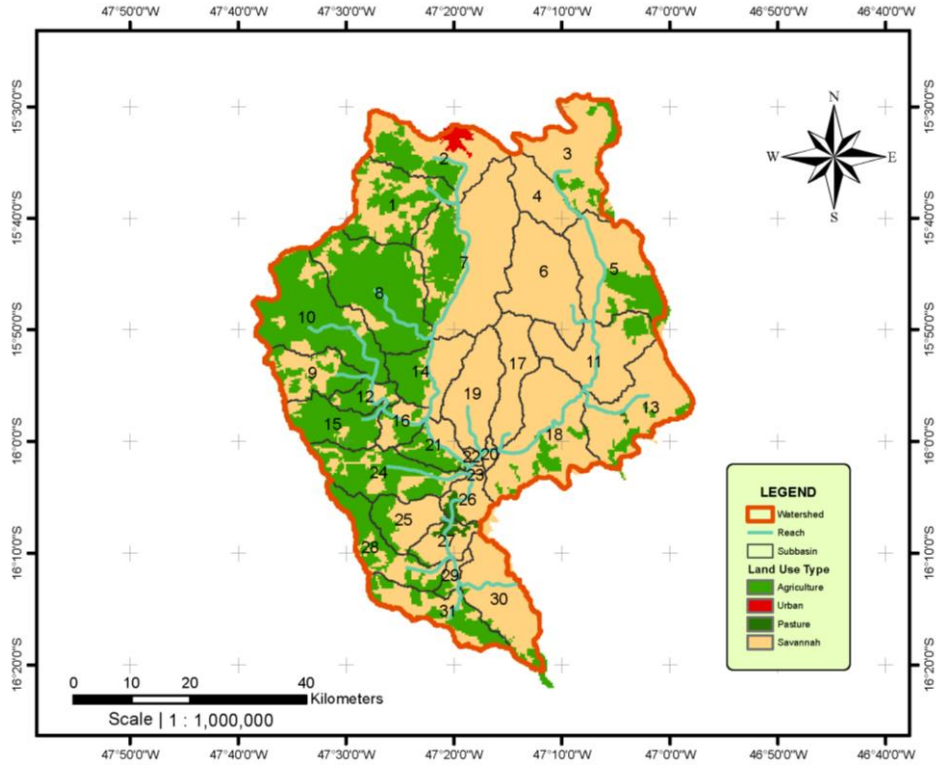


Figure 18. Land-use map of 1984 of the regional watershed model, containing the Preto. The border around the area indicates the watershed boundary, there are 4 land-use types, sub-basis are indicated as well as the river reach.(Embrapa 2016)

Meteorological data was collected from seven meteorological stations. Figure 19 shows the locations of the selected rain gauges, meteorological stations and the hydrological station. The two arrows indicate that the two rain gauges locations are more to the east, slightly outside the domain of the figure.

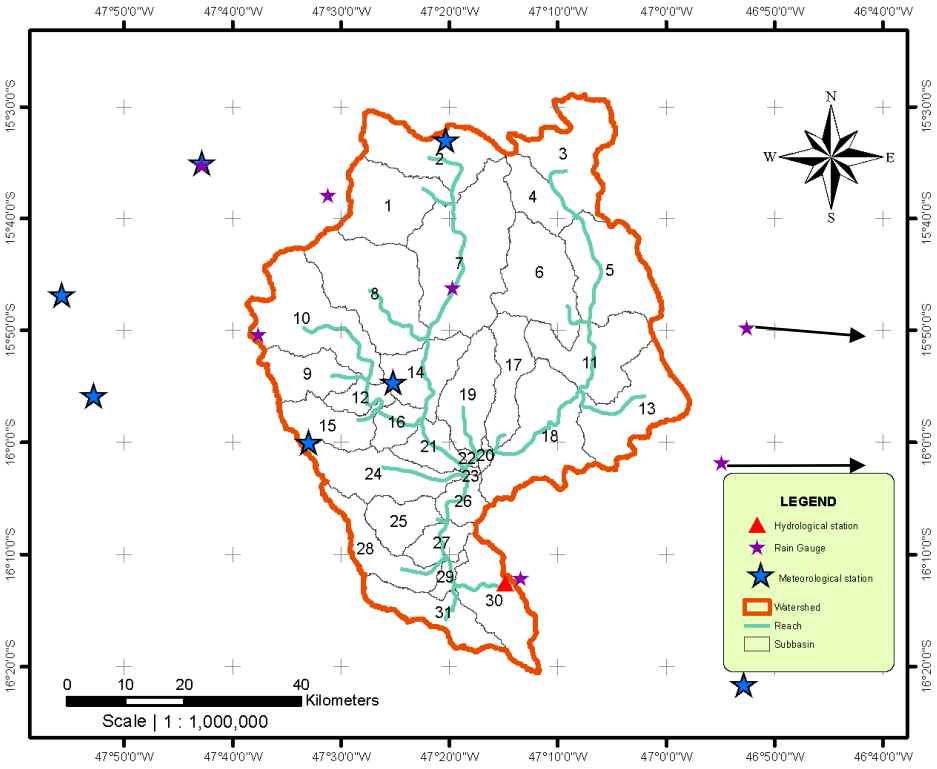


Figure 19. Hydrometeorological stations in the used for the regional watershed model. The red border indicates the watershed boundary, subbasins are indictated. The blue star indicates meteorological stations, the purple star indicates rain gauges and the red trianble indicates a hydrological discharge station Data provided by Embrapa, (2016)

Meteorological data

Meteorological stations were selected based on time series period, percentage missing data (<10%) and availability of data of all the necessary inputs for SWAT; wind speed (m/s), solar radiation (MJ m2 d-1), maximum and minimum temperature (˚C). Figure 20 shows average monthly values for all meteorological input data over the time period 1986 and 2014. Wind speed is highest during the July and October. The average minimum and maximum temperature are respectively 15.1 °C and 26.1°C. Humidity decreases starting in April, when the dry season starts and starts to increase in September as the wet period starts again. During the months April until July solar radiation is the lowest and during the months January, February, September and October it is the highest. The upper right temperature graph shows the same trend as the lower right solar radiation graph.

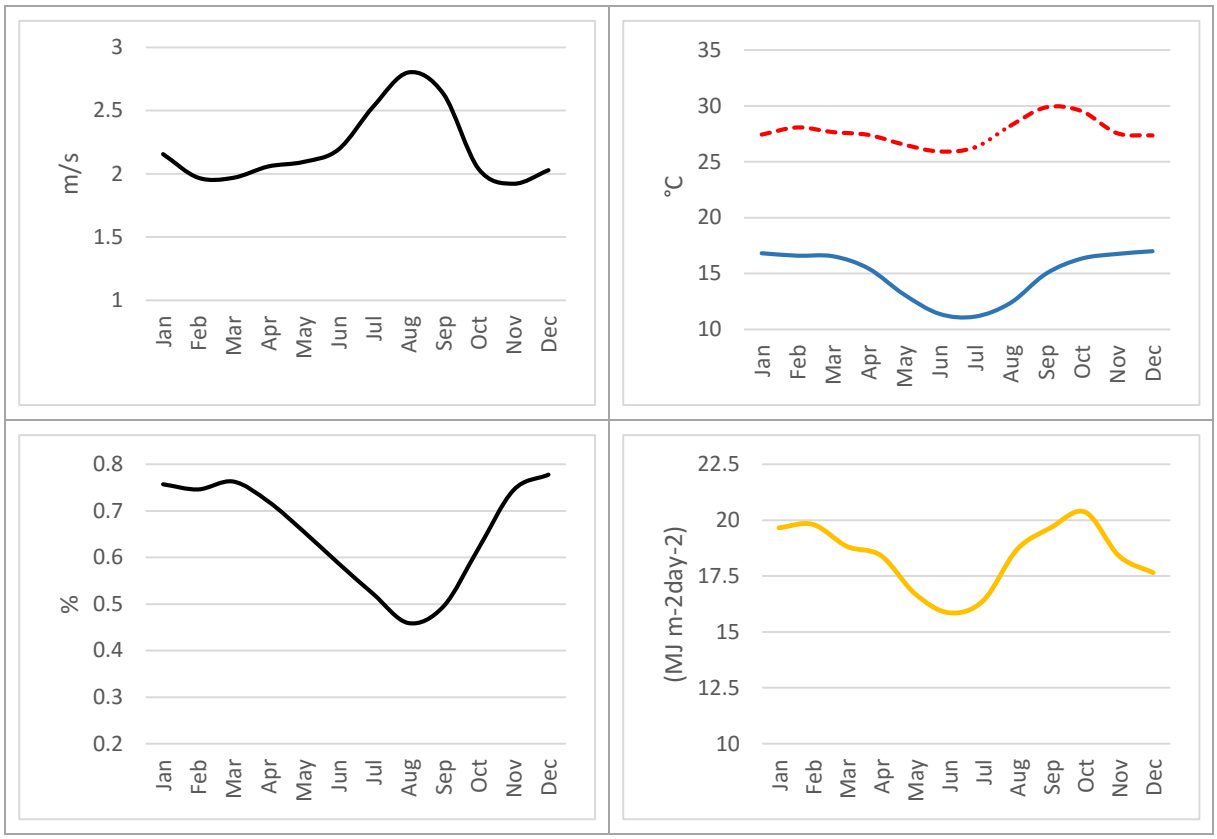


*Figure 20. Four graphs showing average monthly wind speed (m/s)(**A**), average monthly maximum and minimum temperature (°C)(**B**), average monthly relative humidity (%)(**C**) and average solar radiation (MJ/m^2/d)(**D**). Values are averaged over the period 1984 until 2014.*

Precipitation data

Seven rain gauges were selected in and around the regional watershed based on the criteria; missing data, time series period and analysis and comparison between the various stations. During this research the SWAT weather generator was not used to gap fill precipitation data and therefore the selection criteria considering missing data was more strict than for meteorological data (<7%). Figure 21 shows monthly average precipitation between 1984 and 2014 in mm. It clearly shows the typical tropical climate with a wet season from October until April and a dry season between May and September.

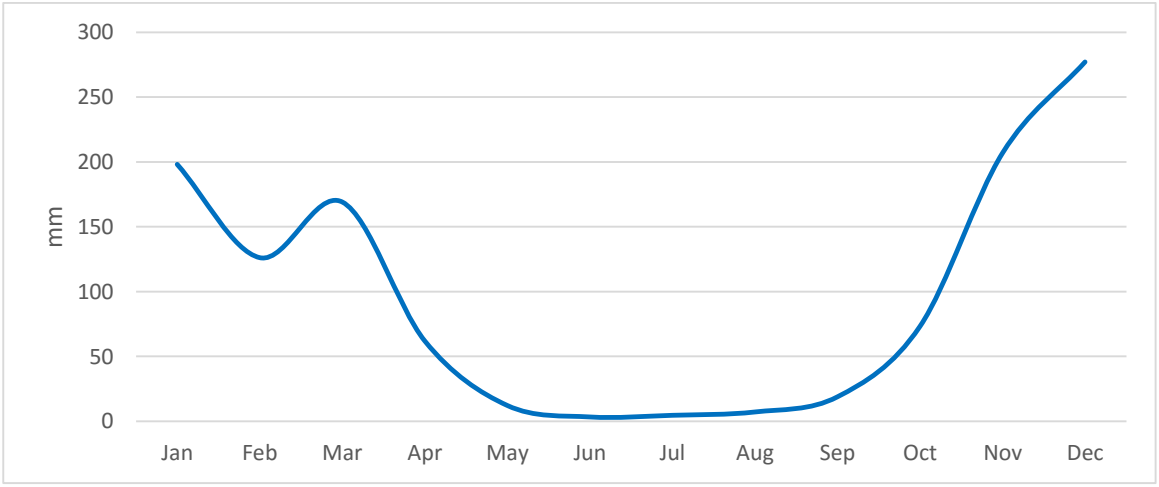


Figure 21. Monthly average precipitation in the regional watershed model (mm/month)

Hydrological data

The red triangle in figure 19 shows the Fazenda Limeira hydrological station, of which discharge data was collected between 1984 and 2013. It contained a time series with average daily values in m3 s-1. The hydrological station is located in subbasin 30, which contains the outlet for the entire watershed. Figure 22 shows monthly average discharge over the selected calibration and validation period from 1984 until 1994 for the regional watershed model. Discharge in the years after this period is influenced by land-use change, the hydropower dam and potentially climate change. The monthly average values in figure 22 thus show a more natural discharge.

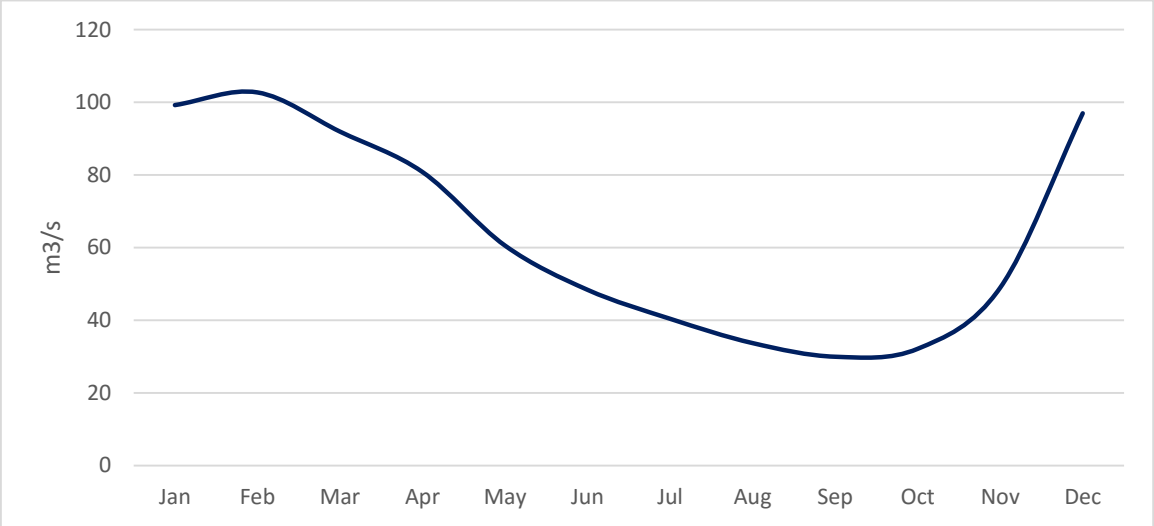


Figure 22. Average monthly discharge measured at the Fazenda Limeira hydrological station during the period 1984 until 1994

Figure 23 shows the monthly average discharge over the period 1984 until 1994. The second half of 1985 is missing. 1990 and 1992 are wet years and 1987 can be considered as a dry year. This has been taken into account while selecting calibration and validation periods for the regional watershed model.

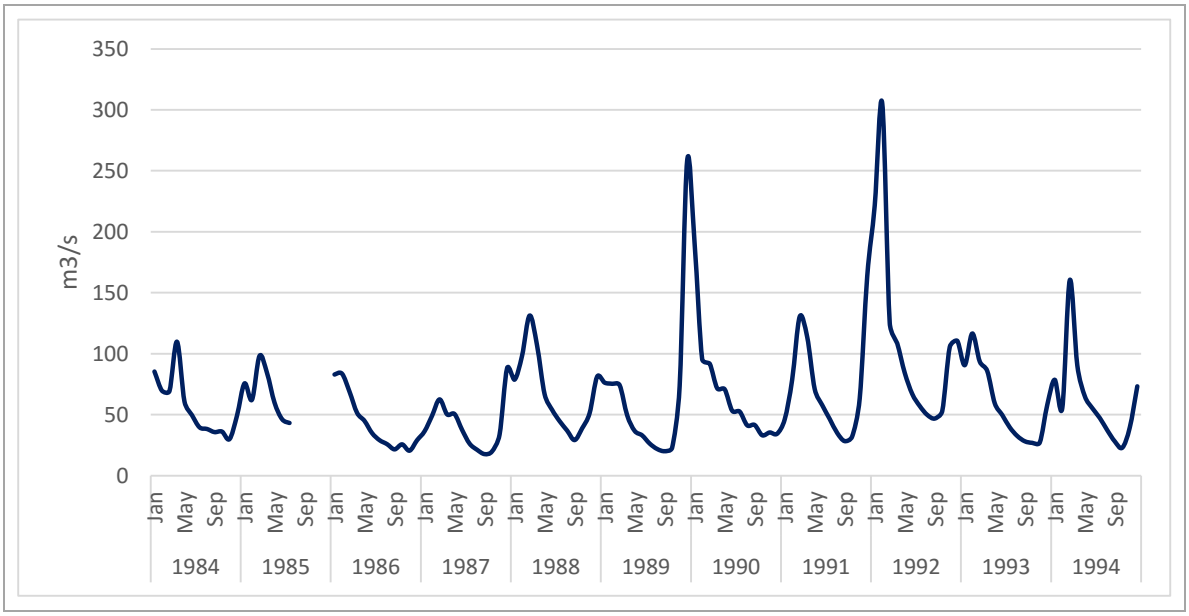


Figure 23. Monthly average discharge at the Fazenda Limeira hydrological station per year (1984-1994)

7.2 Local watershed model data

DEM data is the same as for the regional scale model, with a grid resolution of 30 x 30 meters. Figure 24 shows the elevation map, where the elevation is separated in 10 levels. The maximum elevation difference is 123 meters. Elevation decreases from south towards the north east due to river erosion.

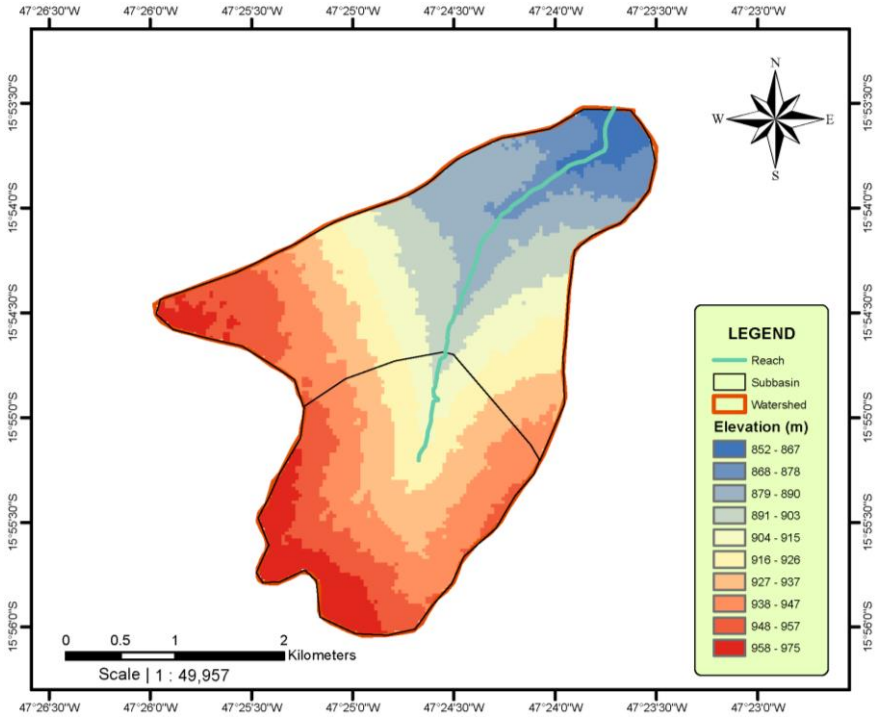


Figure 24. Buriti Vermelho Digital Elevation Map. The red line is the watershed boundary. The elevation map is divided in ten categories. The blue line is the Buriti Vermelho River. Map constructed based on the SRTM DEM 30m (NASA, 2013)

The Buriti Vermelho soil map was provided by the Embrapa (Embrapa, 2016) and is shown in figure 25. The main soil types are latosols, with a small patch of cambisol in the south and gleysol along the river.

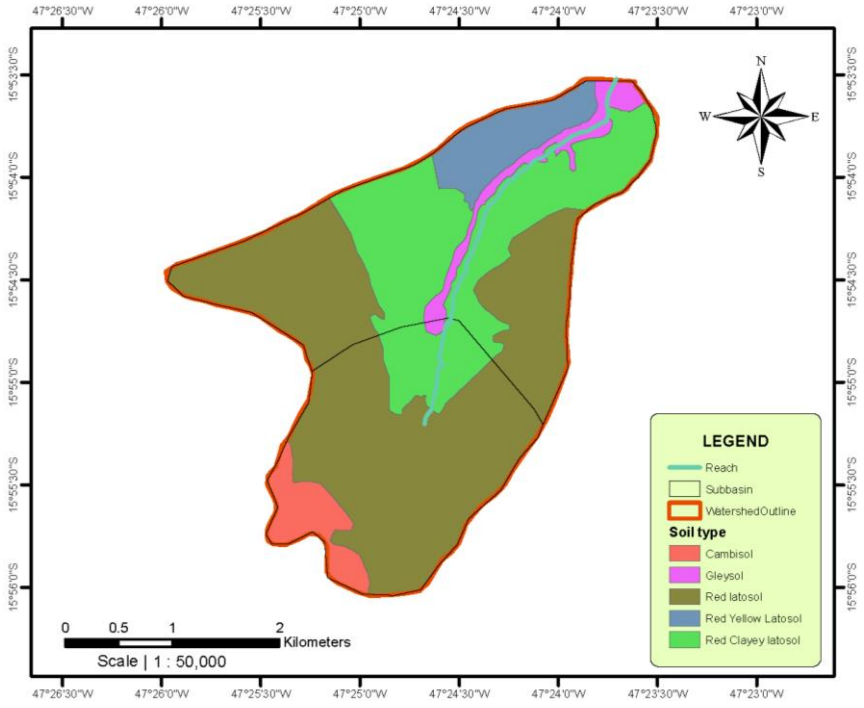


Figure 25. Local watershed soil map. The red line is the watershed boundary. The soil map contains five soil types.. The blue line is the Buriti Vermelho river Data to create map provided by (Embrapa, 2016)

The land-use map data of 2014 was provided by the Embrapa research institute (Embrapa, 2016), shown in figure 26. Rain-fed agriculture crops are mainly soybean and corn. Irrigated agriculture is mainly beans and corn (van Vliet, 2012). The three lower light green circles are pivots, which collect irrigation water from a source outside the local watershed (Rodrigues, 2016). The upper pivot in the downstream area collects water from the Buriti Vermelho river. The patch of land in the North West is filled with small agricultural fields owned by different farmers. These farmers use water from the Buriti Vermelho river to irrigate their fields. The water is collected using irrigation canals that are indicated in figure 27.

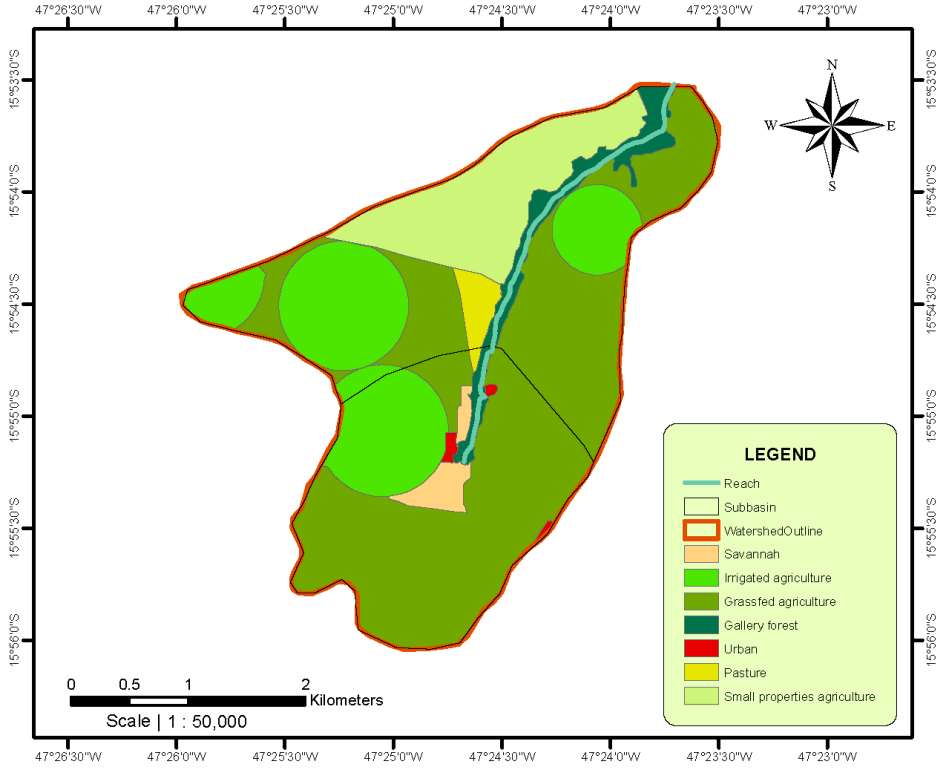


Figure 26. Local watershed land-use map. The red line is the watershed boundary. The land-use map contains five soil types indicated in the legend.. The blue line is the Buriti Vermelho river. Data to create the map provided by (Embrapa, 2016)

Meteorological data was collected from the same stations as the stations for the regional model. The Campbell meteorological station is located inside the local watershed. Figure 27 shows the locations of the selected rain gauges, meteorological stations, hydrological stations, canals and small reservoirs. The PADF and CPAC meteorological stations are indicated. Arrows show the direction and a distance value indicates the location from their symbol. Figure 20 shows the weather data used for the local watershed model, which is the same as for the regional watershed model. SWAT interpolates between the meteorological stations. The Campbell meteorological station will have the highest interpolation weight, because it is located inside the local watershed. The canals drain two reservoirs and bring water to the small properties agriculture land-use type, indicated in figure 27.

The Campbell rain gauge provided precipitation time series from September 2007 until December 2015 and contained precipitation data in millimeters per day. Data collected from the Ana Lado rain gauge contained a time series of millimeters per minute for the period January 2011 until June 2013. Data collected from the Vertedor rain gauge was in units of millimeters per five minutes and ran from February 2014 until November 2015.

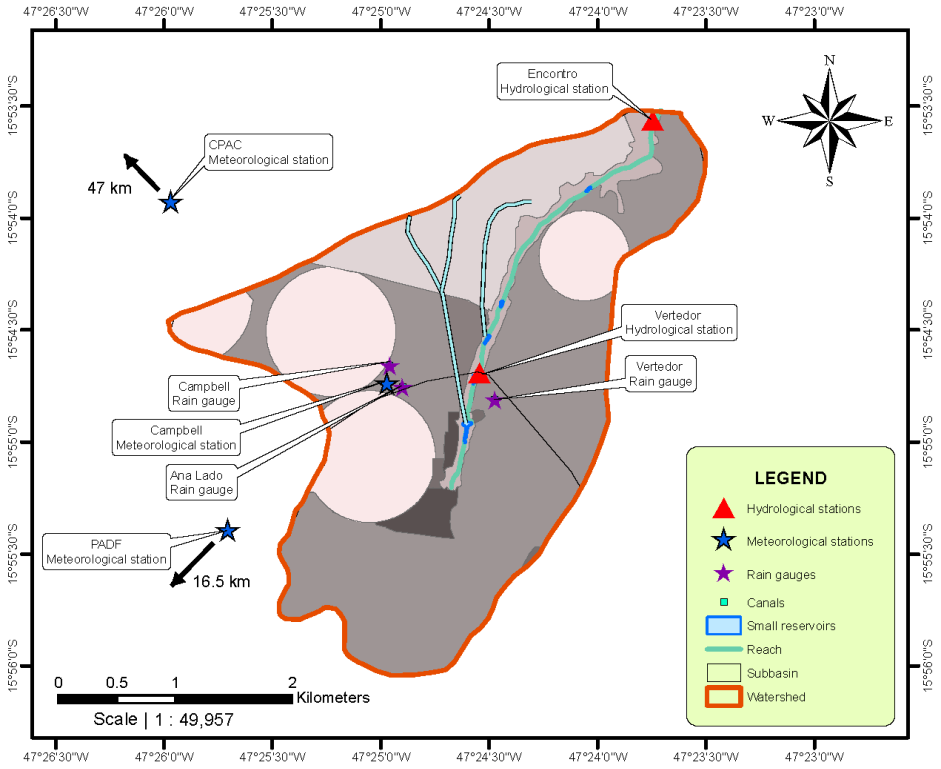


Figure 27. Hydro-meteorological station locations, canals and reservoirs. The red line is the watershed boundary. The blue line is the Buriti Vermelho river. Data to create the map provided by (Embrapa, 2016)

Figure 28 shows the monthly sum of the time series of the rain gauges inside the Buriti Vermelho watershed. During wet periods the monthly sum of measured precipitation can reach 250 mm. The collected data from the Campbell station differs significantly from earlier years and from the Ana Lado measurements from September 2011 until the end of the time series.

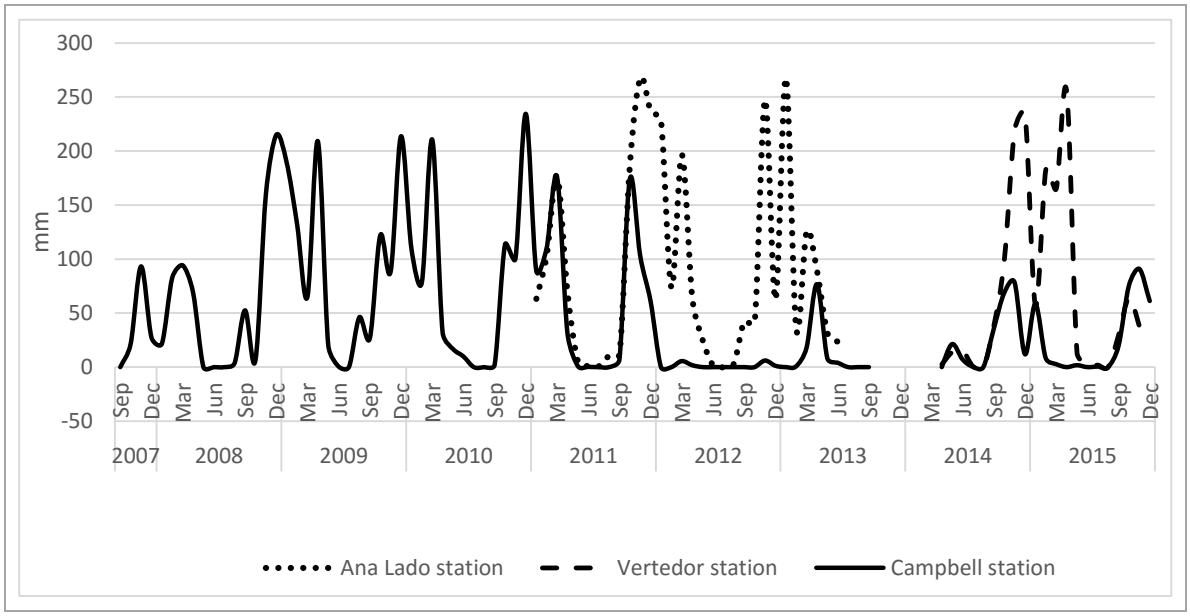


Figure 28. Monthly precipitation sum collected from the rain gauges inside the local watershed. The black line is the time series showing precipitation data from the Campbell station, the dotted line indicates data from the Ana Lado station and the striped line indicates data collected from the Vertedor station.(Embrapa, 2016)

Hydrological data

The Vertedor and Encontro hydrological station provided hydraulic head time series. These were converted to discharge time series which are shown in figure 29. The Encontro hydrological station is located at the watershed outlet and the collected time series was used to validate the downscaling methodology. The Vertedor discharge time series was used to check the quality of the Encontro discharge time series. Moreover the Vertedor hydraulic head time series was used to calculate water being extracted by the irrigation canals. The period between October 2013 and July 2014 does not contain data for the Encontro discharge time series. A trend line is added to figure 29 to show that the Buriti Vermelho river discharges less water, from an average of 0.12 m3/s in 2006 to 0.06 m3/s in 2015. An average 50% decrease.

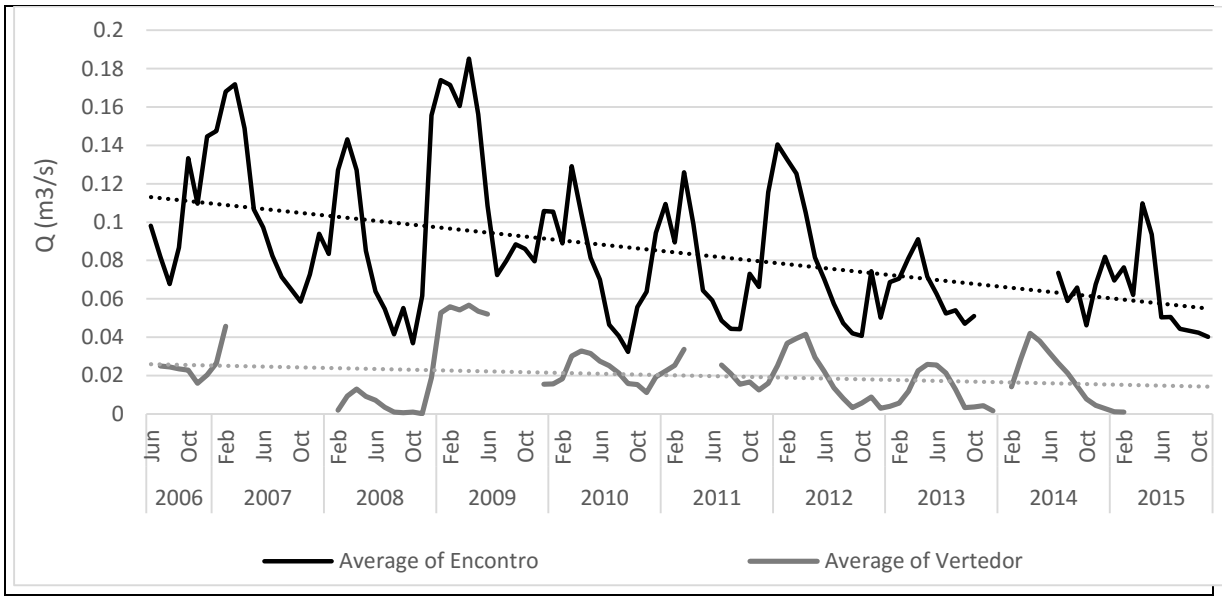


Figure 29. Monthly average discharge time series collected from the Encontro and Vertedor hydrological stations. A linear trend line is added to both time series. Data provided by Embrapa, (2016)

8. Data analysis and preparation

Methods used to analyze, transform and prepare data necessary for the construction of the SWAT models are described in this paragraph. Statistical data analysis and gap filling techniques are also explained, as well as methods used to alter geographical datasets.

8.1 Watershed and HRU data

Watershed delineation and subwatershed discretization

A watershed boundary of the local watershed model was defined by Embrapa based on local elevation measurements. The regional watershed boundaries were provided by Embrapa and based on a digital elevation map (Embrapa, 2016).

Sub-watersheds for the local watershed model were defined by Embrapa and could be imported into the SWAT model as a shape file (Embrapa, 2016). For the regional watershed a tool was used provided by the SWAT program which determines the sub-watershed discretization. For this the tool needs a mask of the watershed area, a shape file of the stream network and the digital elevation map. Each individual tributary of the main river attains a sub-watershed after the SWAT tool operation (Neitsch et al., 2009).

Slope data

The SWAT model base can convert a DEM to a slope map after defining slope groups. In order to define these slope groups a slope map was created from the DEM using the GIS slope tool (spatial analyst), with the setting ‘percent rise’. After analyzing the percentage distribution of slopes from the Symbology tab of the created slope map, three slope groups were defined: under 4.5%, 4.5%-12.5%, and above 12.5%. Using this method gives a more representative distribution of slope classes than randomly defining slope classes. The

resulting slope map is shown in figure 30. The region mostly contains slopes between 0 and 4.5%. Only in the most southern point there is a large patch of slopes above 12.5%. Slightly in front of this region the hydropower dam is built, which profits from this sudden decrease in elevation.

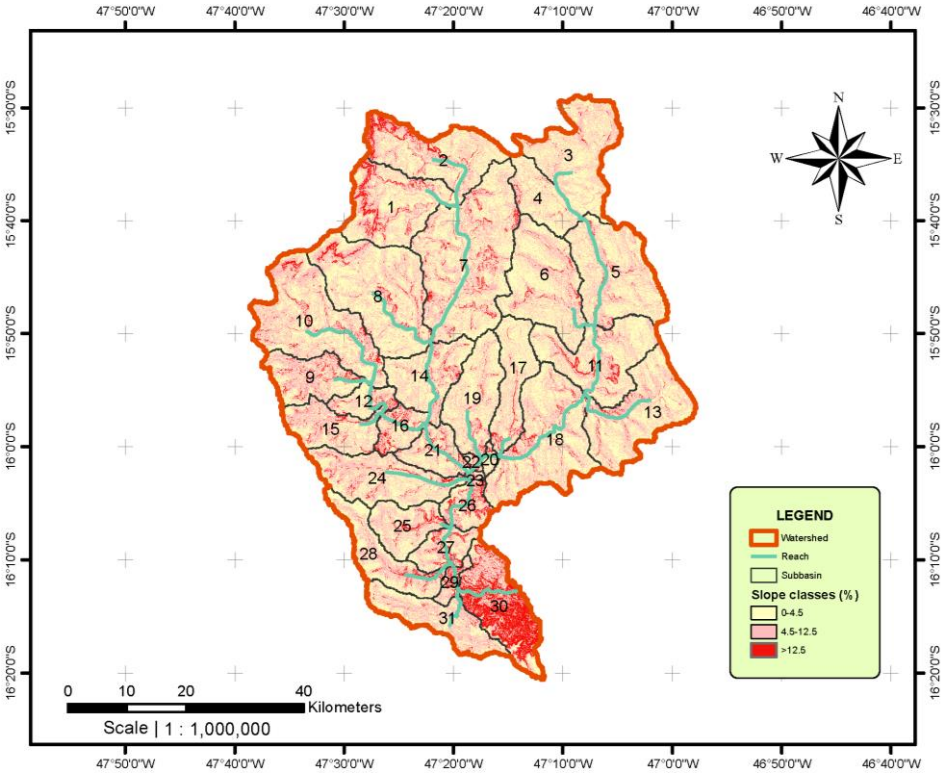


Figure 30. Regional watershed slope map created using the GIS slope tool. The slope map is divided into three slope groups, yellow being nearly flat, pink slightly angled and red are steep slopes. The red line around the watershed is the watershed boundary. The blue line indicates the rivers.

Soil data

The soil maps were provided by the Embrapa research institute for both the regional and the local watershed (Embrapa, 2016). Important soil parameters are soil layer depth (SOL_Z), saturated hydraulic conductivity (SOL_K), rooting depth (SOL_ZMX), the soil available water content (SOL_AWC) and the percentage distribution of clay, silt and sand. The most common soil types are variations of latosols and cambisol. Table 6 shows values for these important parameters. Both soil types have two soil depth layers, with differing parameter values as can be seen in table 7. Latosol has a very high percentage of sand and thus also a very high saturated hydraulic conductivity compared to Cambisol. The depth of the latosol soil type also has a higher value.

Table 6. SWAT soil input parameters, their units and values for Latosol and Cambisol

Parameter	*Unit*	*Latosol*	*Cambisol*
SOL_Z1	mm	100	290
SOL_Z2	mm	3100	2090
SOL_AWC	mm	0.13	0.24
SOL_ZMX	mm	3100	2090
SOL_K1	mm/hr	±1600	307
SOL_K2	mm/hr	±1100	82
Clay	%	8	34
Silt	%	9	32
Sand	%	83	34

Land-use data

A land-use map was provided by Embrapa based on land-sat TM5 images. The data only provided a description of the land-use types and not the required parameter values. The curve number value is an

example of an important input parameter that is determined by the land-use map. Inside the SWAT land-use database, land-use types are specified and can be selected (Arnold, Kiniry, et al., 2012). The land-use map of the regional watershed model contains 4 land-use types; agriculture, pasture, savannah and urban. The local watershed model also has information about whether the agricultural land is irrigated or rainfed agriculture. Van Vliet (2012) described the most commonly used crops for both land-use types in the Buriti Vermelho watershed. For irrigated agriculture soybean used 50% of the area and corn 41% and for rainfed agriculture beans used 36% of the area and corn 28% (van Vliet, 2012). Irrigated and rainfed agricultural area of the land-use map in the local watershed model, was subdivided according to the crop distribution by van Vliet.

8.2 Modelling irrigation

The local watershed model contains two large irrigation pivots that collect water from a source outside the watershed boundary. Irrigation of these pivots will, especially during the dry season, impact the hydrological balance and are taken into account during modelling of the local watershed model. A third pivot in the local watershed collects water from the Buriti Vermelho river. As the extracted water is applied back inside the watershed area, this pivot will not be modelled as an irrigated area and will rather be seen as regular cropland. This assumption is plausible because the output generated by the SWAT model are monthly values.

Irrigation is not incorporated in the regional watershed model. Because of the large area of this watershed the ratio between irrigated agricultural fields near the border and in the center is very small. Most irrigated areas use water from inside the watershed and not from an outside source. It is thus assumed that this small portion of irrigated fields along the watershed border that do use water from an outside source have a minimal influence on the entire watershed. Moreover information is not available about the sources of water of the irrigated fields.

To incorporate irrigation in the SWAT model an auto-application of irrigation can be selected. This application of water can be triggered by a water stress threshold or a soil water deficit threshold (Neitsch et al., 2009). If plant stress falls below this water threshold due to water stress, the model will apply water to the area. The user is required to enter this maximum value of available water for irrigation per application. The selected value was based on the daily water need of the standard grass crop, called the 'reference crop evapotranspiration' based on different climatic zones *(Brouwer,et al., 1986)*. Table 7 shows how the 'reference crop evapotranspiration' is dependent on climatic zone and mean daily temperature.

Irrigation is applied mainly during the dry season. In this season temperatures are above 25°C and as there is hardly any precipitation in this period the climatic zone is considered as semi-arid. The maximum daily irrigation amount was chosen to be 9 mm/day based on table 8.

Table 7. Average daily water need of the standard grass during irrigation season (Brouwer, Heibloem, & Division, 1986)

Climatic zone	Mean daily temperature		
	Low	Medium	High
	(less than 15°C)	(15-25°C)	(More than 25°C)
Desert/arid	4-6	7-8	9-10
Semi arid	4-5	6-7	8-9
Sub-humid	3-4	5-6	7-8
Humid	1-2	3-4	5-6

8.3 Time series analysis and adjustments

The collected datasets for discharge, precipitation and weather data consisted of time series, which required various statistical tools to validate the quality of the data. Outliers were identified and excluded from the time-series using the z-test. Time series were validated by comparing time series from various stations with each other to determine the R^2 correlation. Missing data in meteorological stations and rain gauges were gap filled using linear regression with nearest station by R/d ratio (LRNR/d) technique. An attempt was made to estimate discharge in missing periods, such as was the case for discharge in the Encontro hydrological station time series, using watershed area and an upstream hydrological station. Units often had to be converted to the specific unit required as input for SWAT. And finally using a canal discharge model, average daily water extraction (m^3/d) by two canals in the Buriti Vermelho river was calculated and summed with the daily observed discharge time series observed in the Encontro hydrological station.

This subparagraph explains the statistical analyses methods and gap filling methods used to prepare the SWAT input data and it presents examples of the application of the techniques with the collected data.

8.3.1 Input unit conversion

SWAT weather data

SWAT mainly needs meteorological data to calculate evapotranspiration (ET) using either the Priestley Taylor equation, Hargreaves equation or the Penman Monteith equation. Because of the more accurate results in this research the Penman Monteith equation is used:

$$\lambda E = \frac{\Delta(R_n - G) + \rho_{air} c_p \frac{(e_z^0 - e_z)}{r_a}}{\Delta + \gamma\left(1 + \frac{r_c}{r_s}\right)} \quad (1)$$

where

λE	*latent heat flux density*	*(MJ m^{-2} d^{-1})*
R_n	*solar radiation*	*(MJ m^{-2} d^{-1})*
G	*heat flux density to the ground*	*(MJ m^{-2} d^{-1})*
ρ_{air}	*air density*	*(kg m^{-3})*
c_p	*specific heat at constant pressure*	*MJ kg^{-1} ˚C^{-1}*
r_a	*(aerodynamic resistance) (s m-1*	*s m^{-1}*
e_z^0	*saturation vapor pressure of air at height 0*	*kPa*
e_z	*saturation vapor pressure of air at height z*	*kPa*
r_c	*plant canopy resistance*	*s m^{-1}*
Δ	*Slope of the saturation vapor pressur-temperature curve*	*kPa ˚C-1*
γ	*psychrometric constant*	*kPa ˚C^{-1}*

The necessary weather SWAT input data required for this equation is summarized in table 8.

Table 8. SWAT weather input parameters

Weather input parameter	*Abreviation*	*Unit*
Solar radiation	SLR	MJ/m2/day
Wind	WND	m/s
Maximum and minimum temperature	TMP	°C
Relative humidity	HMD	%

The collected weather data contained time series, which needed to be averaged to contain daily averaged values as input for SWAT.

Solar radiation

Solar or shortwave radiation was measured in amount of sun hours per day. This needed to be converted to mega joule per square meter per day. This is the unit for solar radiation determined with the Angstrom formula which relates solar radiation to extraterrestrial radiation and relative sunshine duration:

$$SLR = (a_s + b_s \frac{n}{N}) R_a \quad (2)$$

Where:

SLR	*solar radiation*	*(MJ m^{-2} day^{-1})*
n	*actual duration of sunshine*	*(hour)*
N	*maximum possible duration of sunshine*	*(hour)*
R_a	*extraterrestrial radiation*	*(MJ m^{-2} day^{-1})*
a_s	*regression constant, expressing fraction of extraterrestrial radiation reaching the earth on overcast days*	*(-)*
$a_s + b_s$	*fraction of extraterrestrial radiation reaching the earth on clear days (n = N)*	*(-)*

The recommended values for a_s and b_s are respectively 0.25 and 0.50. N can be determined with the equation:

$$N = \frac{24}{\pi} \omega_s \quad (3)$$

Where ω_s is the sunset hour angle in radians, which can be determined using the following equation:

$$\omega_s = \arccos[-\tan(\varphi)\tan(\delta)] \quad (4)$$

in which φ is latitude, which is given in the data set of the meteorological station and δ is solar decimation. Latitude in radians is calculated using:

$$[radians] = \frac{\pi}{180} [decimal\ degrees] \quad (5)$$

and solar decimation is calculated using:

$$\delta = 0.409 \sin\left(\frac{2\pi}{365} J - 1.39\right) \quad (6)$$

where J is the number of days in the year. Lastly we need the extraterrestrial radiation ($R_{a)}$. Ra for each day of the year and for different latitudes can be estimated from the solar constant, the solar declination and the time of the year by:

$$R_a = \frac{24(60)}{\pi} G_{sc} d_r [\omega_s \sin(\varphi)\sin(\delta) + \cos(\varphi)\cos(\delta)\sin(\omega_s)] \quad (7)$$

Where

R_a	*is extraterrestrial radiation*	*(MJ m^{-2} day^{-1})*
G_{sc}	*is solar constant*	*(0.0820 MJ m^{-2} min^{-1})*
d_r	*is Inverse relative distance Earth-Sun*	*(-)*

To determine the inverse relative distance Earth-Sun:

$$d_r = 1 + 0.033\ cos\left(\frac{2\pi}{365} J\right) \quad (8)$$

Using these equations in Excel resulted in the calculation of the necessary input time series of solar radiation for SWAT.

The other input time series in table 8, i.e. relative humidity, wind speed and minimum and maximum temperature required a unit that did not need a complex conversion from the observed data. Hourly or five minute measurements had to be averaged to daily values, but nothing more complex than that.

Data quality testing

Outliers were detected using the z-test, for both the weather parameters relative humidity, solar radiation and temperature. Precipitation and wind were not tested as they have a high spatial and temporal variability. The z-test is a statistical test that gives a z-score to each data point based on the normal distribution of the entire population (van Vliet, 2012). First calculate the standard error of the mean, using;

$$z.score = \frac{Sample - \mu}{\sigma} \tag{9}$$

In which μ is the population mean and σ is the standard deviation. If the z-score is higher than 3 the sample value is considered as an outlier and is removed from the time series.

The precipitation time series for the local watershed model is shown in figure 28. It shows how three rain gauges were used. The longest time series was collected from the Campbell rain gauge. From figure 28 we can see that from December 2011 onwards the precipitation data seems flawed. It underestimates the precipitation measurements compared to earlier years. The Ana Lado precipitation data confirms this suspicion, as the time series starts in 2011 and overlaps with the Campbell time series, revealing a large difference in observed precipitation values. The Ana Lado observed values are much higher. This resulted in the choice to use the Campbell precipitation time series until December 2011, then adding the Ana Lado time series, which ends in June 2013.

Gap filling

Gap filling is an essential tool when analyzing and preparing time series as input for hydrological models. Often time series contain missing data and after quality analysis of the time series often outliers are removed from the time series. These removed outliers leave gaps in the time series. There are various methods of gap filling which can be categorized as deterministic methods, stochastic methods and artificial intelligence methods (Campozano, et al., 2014). Deterministic methods always produce the same output given initial conditions. Examples are the arithmetic mean of the corresponding values of stations near the station location of the time series that is missing data. Inverse distance weighting is probably the most commonly used method to estimate missing data in hydrology and geographical sciences (Campozano, et al., 2014). Xia, et al. (1999) examined various methods for gap filling; arithmetic averaging, inverse distance interpolation, normal ratio method, multiple regression analysis and universal kriging. Xia concluded that multiple regression and kriging produced the best results. A second gap filling category is stochastic methods, these provide a probabilistic estimate. These methods generally outperform deterministic methods, but are computationally and thus time demanding (Campozano et al., 2014). An example of stochastic gap filling is artificial neural network (ANN). The last gap filling method category is artificial intelligence methods, which use complex mathematical formulation and require computation with high computational cost. These methods are effective when dealing with non-linear relationships. An example is the Monte Carlo Markov Chain algorithm (EM-MCMC), which is based on expectation-maximization (Campozano et al., 2014).

Rain data gap filling regional watershed model

In this research gap filling was done using the deterministic method; multiple regression analysis, which was selected by Xia et al. (1999) as the best performing deterministic method next to universal kriging. Gap filling is only applied for data collected from the rain gauges. Figure 31 shows the used rain gauge time series and in which periods data is missing as a fraction of the total daily time series data between 1984 and 2015. From the figure can be seen that there are no time series used with more than 7 percent of missing data. This was one of the selection criteria in determining which precipitation stations would be used. Missing data between 1994 and 2007 is not counted, as these periods are not used for analyses, calibration or validation, resulting in a straight line during these two years for all the time series.

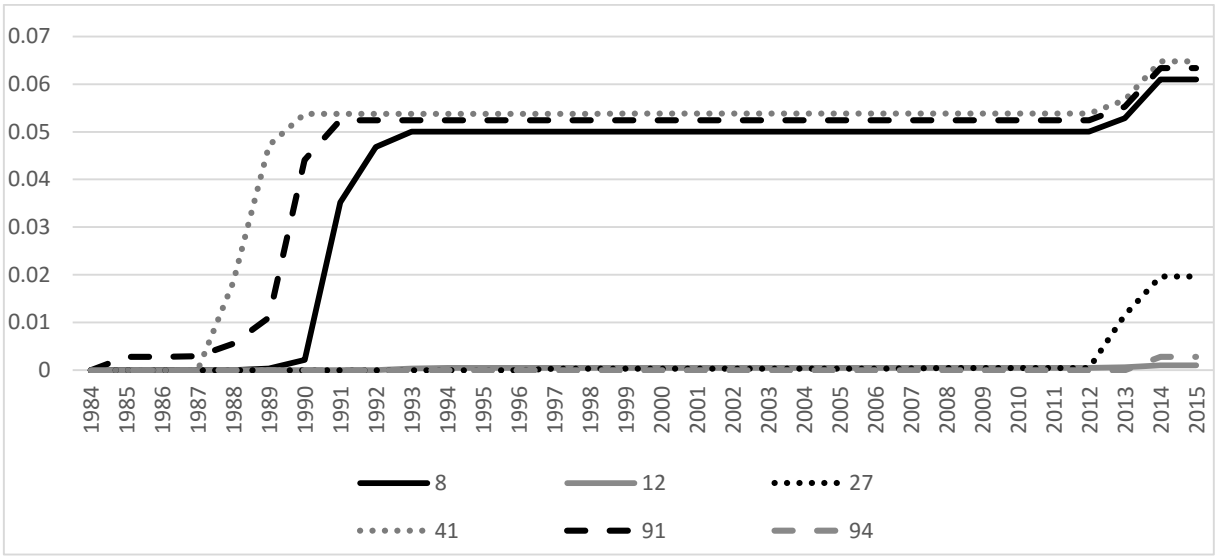


Figure 31. Cumulative percentage of missing data gaps for the selected precipitation stations

The multiple regression method used is called linear regression with the nearest station by R/d ratio (LRNR/d) (Campozano et al., 2014). This method requires to first use;

$$V_{est} = b * V_i + a \tag{10}$$

in which V_i is the value of a near station, V_{est} is the estimated value and a and b are regression parameters. After this use:

$$V_{est} = \frac{\sum_{i=1}^{n} w_i^2 * (b * V_i + a_i)}{\sum_{i=1}^{n} w_i^2} \tag{11}$$

in which w_i is the weight of the i-th precipitation station, which depends on the ratio between distance d and the regression value R. (w_i = R/d). Per precipitation station three surrounding rain station were selected, the distances were measured in the GIS environment and R regression values were calculated over the entire time series length. In Excel the equations were used to determine V_{est} time series for each precipitation station. The gaps in the observed precipitation time series were then filled using the corresponding V_{est} time series. Figure 32 shows part of a precipitation time series that was gap filled using this method. The thick black line indicates the time series that needs gap filling. A period of data is missing between December 3rd 1986 and the 2nd of January 1987. The grey lines show the time series of the nearby rain gauge stations and the dotted line shows the result of the LRNR/d gap filling method.

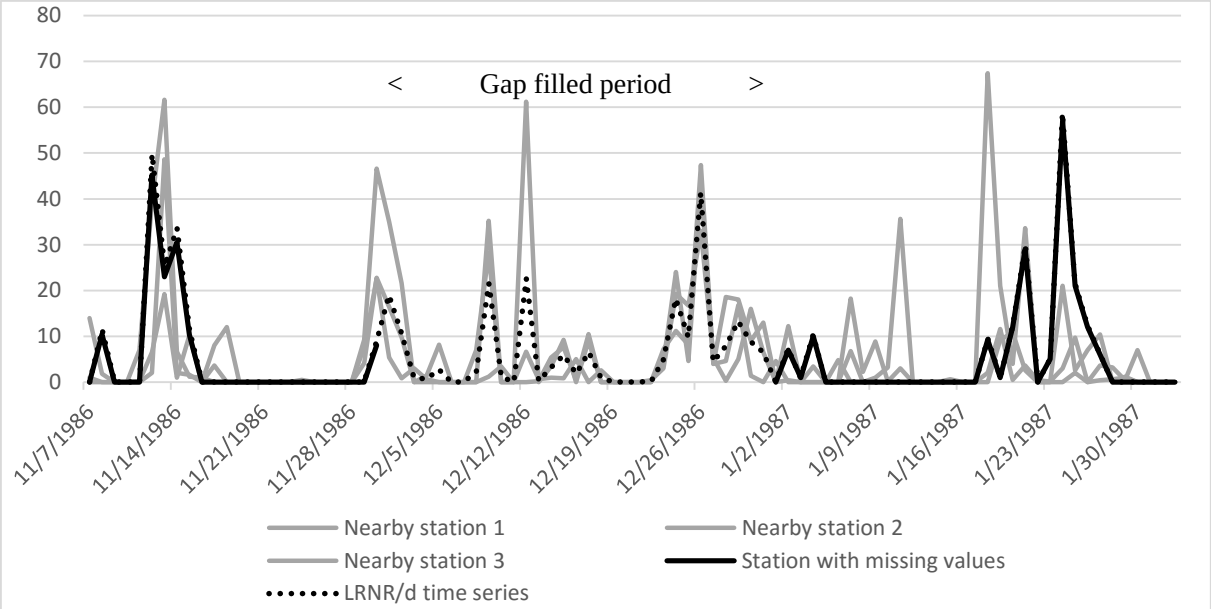


Figure 32. Example showing result of LRNR/d gap filling method. The black line is the time series of the station with missing values. The grey lines are nearby stations, with each a different distance from the station with missing values and the black dotted line is the resulting time series after application of the LRNR/d method for gap filling

Rainfall time series construction local watershed model

Because of the small area size of the watershed (±9.3km^2) surrounding rain gauges outside of the watershed were not used. The reason for this is the heterogeneous tropical climate, which is known for high spatial variability in precipitation events. Moreover the nearest precipitation station outside the local watersheds was too far away. Gap filling was thus not possible. However after analyzing precipitation-discharge graphs made from the Campbell rain gauge collected precipitation data and the Encontro station discharge data, it was concluded that the precipitation data after 2011. This can also be concluded from figure 33, which shows precipitation data from the Ana Lado rain gauge and the Vertedor rain gauge. The Campbell precipitation time series was thus only used from 9/2007 until 6/2011 then the Ana Lado time series was used until 6/2013, then there is a gap between 6/2013 and 3/2014, from this point there is precipitation data from the Vertedor rain gauge until 9/2015.

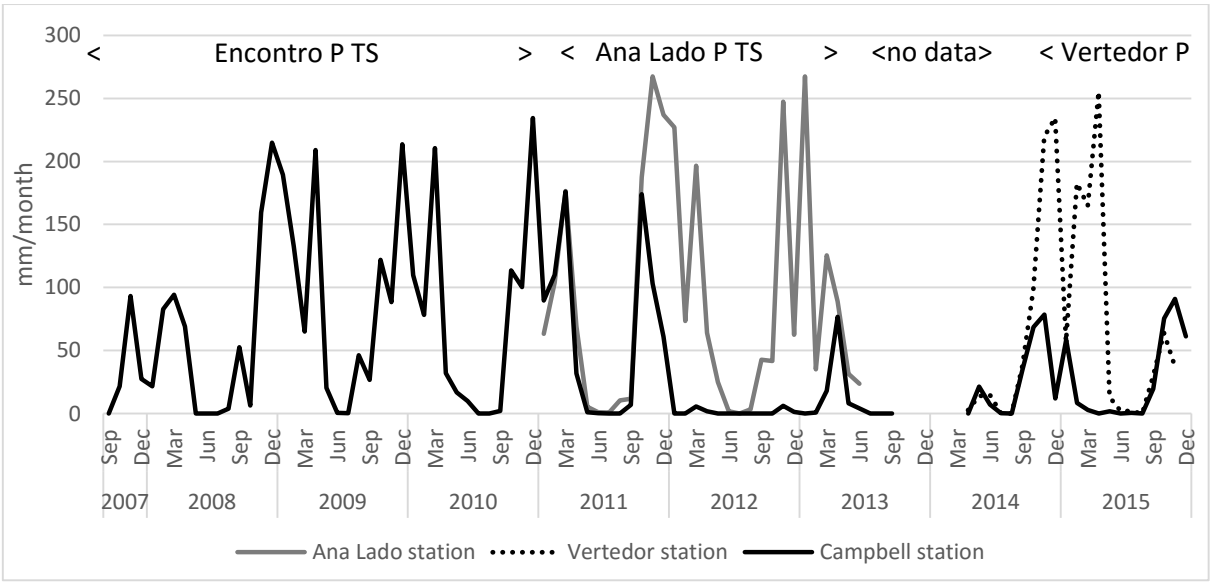


Figure 33. Monthly average precipitation measurements from the Ana Lado, Vertedor and Campbell rain gauge. The text above the chart describes which parts of each precipitation (P) time series (TS) was used as data in the local watershed SWAT model. Between 2013 and 2014 there is a period of no data.

Discharge gap filling local watershed model

An attempt was made to gap fill the missing data period of the discharge time series collected from the Encontro hydrological station, using time series data from the upstream Vertedor hydrological station. The ratio between the subbasin sizes was used to create a time series that can be used to replace missing data points, according to:

$$Q_{encontro} = \frac{A_{Vertedor}}{A_{Encontro}} * Q_{vertedor}$$

in which $Q_{encontro}$ and $Q_{vertedor}$ are the discharge time series measured respectively at the Encontro and Vertedor hydrological station. *A* is the sub-watershed area size. The resulting discharge time series is Q_{gap}. The results of this gapfilling method did not produce realistic looking results and were not used to gapfill the discharge time series of Encontro, which contained a large gap in 2013.

Weather generator SWAT

The SWAT model allows users to use time series input which has missing data. The user is of course expected to keep this percentage to a minimum. The tool inside the SWAT model is called the WXGEN weather generator model (Sharpley and Williams, 1990) and fills missing data using data available in surrounding meteorological stations or rain gauges. In this research the precipitation time series are gap filled manually, but the meteorological time series for solar radiation, wind speed, relative humidity and maximum and minimum temperature are gap filled using the WXGEN weather generator model. The required input for the weather generator are average monthly values and their standard deviations for all the time series that need to be gap filled. The monthly averages and standard deviations are determined using the entire time series length. To acquire these averages and standard deviations the pivot tool in Excel was used, which allows users to group a time series into months and select a formula that does an operation with all the time series data in that group. Using this method for example all the daily wind speed data points in the group 'January' can be averaged or the standard deviation can be calculated.

Next to standard deviations and averages the average monthly dewpoint temperature needed to be calculated. This was done using a dewpoint calculation program by Liersch (2003), which is based on Allen, R.G. (1998). The program firstly calculates saturation vapour pressure e_s from daily minimum and maximum temperature (equation 12). Then the average daily actual vapour pressur e_a is calculated using saturation vapour pressure e_s and daily average relative humidity RH (equation 13). According to Allen,

$$e_s = 0.6108 * exp^{[(17.27*T)/(T+237.3)]} \quad (12)$$

The unit for saturation vapour pressure is then converted from kPa to mbar and can then be used in:

$$e_a = RH * \frac{e_s}{100} \quad (13)$$

The daily dewpoint temperature *dew* is approximated using the following equation:

$$dew = (234.18 * log_{10}(e_a) - 184.2/log_{10}(e_a) \quad (14)$$

in which:

e_s = *saturation vapour pressure [mbar]*
e_a = *actual vapour pressur [mbar]*
exp = *207183 (natural logarithm base)*
T = *air temperature [°C]*
RH = *relative humidity [%]*
Dew = *dewpoint temperature [°C]*

Using minimum and maximum saturation vapour pressure is derived twice using equation 12 ($e_{s\ min}$ and $e_{s\ max}$). The e_a used in equation 13 is the mean of the two values.

The program output gives daily dew values after entering the relative humidity and minimum and maximum time series, based on equation 14. These daily dewpoint temperature are then monthly averaged using the

pivot tool method in Excel as explained earlier. The monthly average values are then used as input for the WXGEN weather generator model in SWAT.

Canal model

Figure 27 in paragraph 7.2 shows two irrigation canals draining water from small reservoirs of the Buriti Vermelho river. The extracted water from the Buriti Vermelho river needs to be calculated and added to the observed discharge of the Vertedor and Encontro hydrological station time series in order to attain natural river discharge time series data. The amount of water extracted by the canals is proportional to the water level and the diameter of the pipe. The two irrigation canals connected to the Buriti Vermelho river thus do not have a discharge control structure. In order to estimate the extracted water amounts of these two canals a program is used which estimates water amounts extracted from river using reservoirs, irrigation channels and without discharge control structures. The methodology and the tool are developed by Rodrigues, L. & Torres, M. (2015).

FIigure 34. Images of the reservoirs in the Buriti Vermelho watershed and one of the two irrigation canals (Rodrigues & Torres, 2015)

Figure 34 shows the two reservoirs (left and middle image) which are drained using an irrigation pipe canal (right image). The extracted water is calculated using water level measurements a short distance downstream of the small reservoir as illustrated in Figure 35 with the red mark. The blue line indicates the river, the triangle indicates the reservoir, the yellow line the irrigation pipe and the red mark is the hydrological measurement station. In this research hydraulic head time series are collected from the Vertedor hydrologic station, shown in figure 27. The model requires water level measurements of one value per five minutes. The model also needs information about the maximum water level in the reservoir and the pipe diameter.

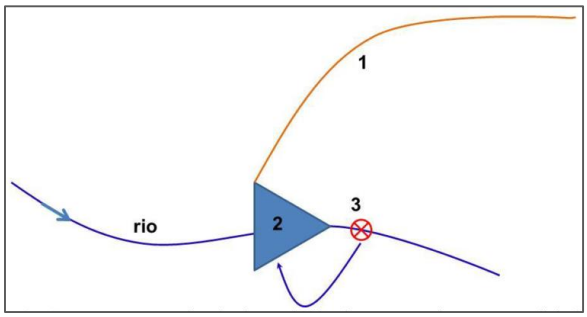


Figure 35. Graphical representation of a river (rio), irrigation canal (1l, a reservoir (2) and a hydrological station (3) (Rodrigues & Torres, 2015)

The model output is the sum of the observed water level input data and the water level change due to extraction by the canals. Using the height and width of the measuring weir, located at the Vertedor hydrological station, as coefficients water level values can be converted to discharge in m^3/s. This results in a discharge time series that represents the natural flow conditions of the Buriti Vermelho river at the Vertedor hydrologic station. The difference between the observed values and the values calculated by the model gives the amount of water extracted by the irrigation canals. These daily differences are summed with the observed discharge values from the Encontro hydrological station to attain daily natural discharge time series. This resulting discharge time series is used to validate the predicted discharge using the downscaling methodology.

Final discharge and precipitation time series

The most important time series for a rainfall runoff watershed model are precipitation as input and discharge for the calibration and validation. Comparing precipitation data and discharge data in the same graph was a method of graphically assessing the quality of the discharge data and of the precipitation data. If in very wet periods no increase in discharge was visible, something was off about the discharge data in this period. The validation of the downscaling methodology was done with the discharge time series collected from the Encontro hydrological station, which was summed with the time series that was calculated using the canal model (Rodrigues & Torres, 2015) to attain a natural discharge time series for the Buriti Vermelho river. This time series is presented in figure 36 together with the precipitation time series data set which is constructed using data from the Campbell rain gauge, Ana Lado rain gauge and the Vertedor rain gauge

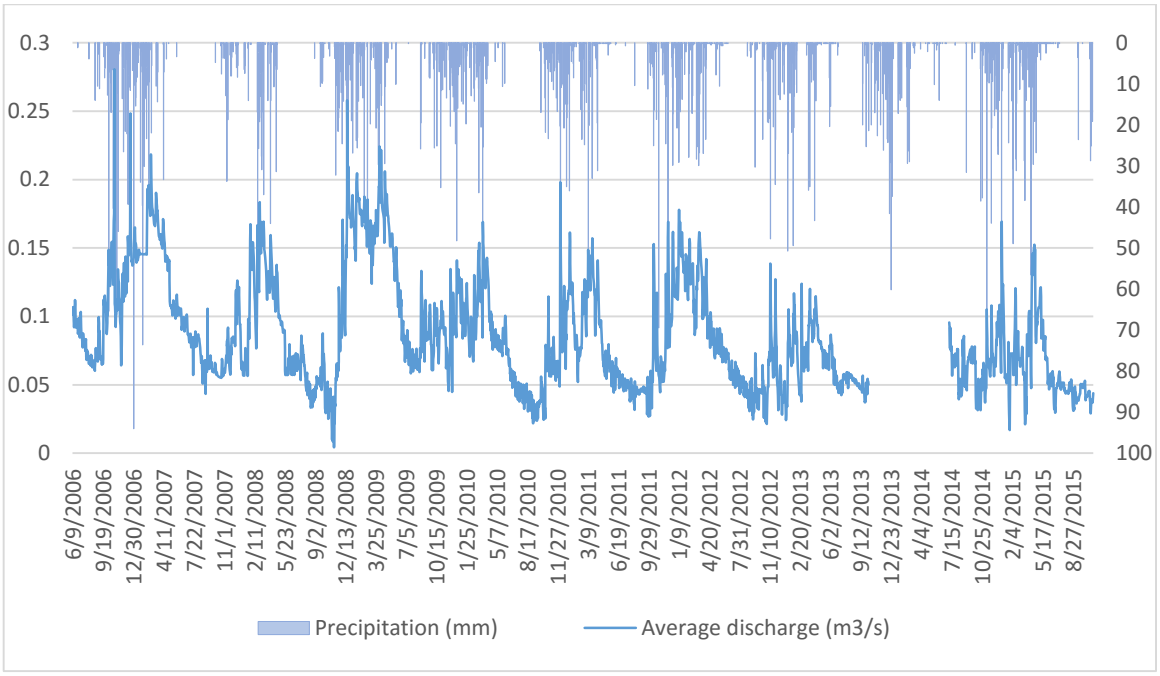


Figure 36. Daily sum precipitation (mm) of the combined time series data set from the Campbell, Ana Lado and Vertedor rain gauges compared to average daily discharge data (m3/s) from the Encontro hydrological station. The data is used in the local watershed model. Precipitation values are presented on the right axis.

9. Sensitivity Analysis

To calibrate physically based hydrologic models it is important to identify sensitive parameters with respect to their impact on model outputs. This can be examined by performing a sensitivity analysis. Determining sensitive parameters can improve understanding of hydrological processes and lead to better estimation of parameter values to reduce uncertainty in the model output (Lenhart et al, 2002). Shen *et al (* 2010) performed First-Order Error Analysis to determine the effect of parameter uncertainty on the SWAT model. Zhang *et al (*2009) did a calibration and uncertainty analysis using Bayesian Model Averaging. Tolson *et al, (*2008) performed and compared two methods of uncertainty analysis and concluded that dynamically dimensional search (DDS) was more effective than the much used generalized likelihood uncertainty estimation (GLUE). There are many different methods and tools for uncertainty analysis and to determine parameter sensitivity.

9.1 Influence coefficient method

In this research the influence coefficient method is used to compute sensitivity coefficients. It is commonly used for hydrological modelling (Helsel *et al,* 1992). The method determines sensitivity by changing independent variables individually, while holding the rest of the parameters constant. The method computes a sensitivity coefficient. This value represents the change of a response variable resulting from change of an explanatory variable:

$$\frac{\Delta F}{\Delta P} = \frac{F(P_1, P_2, \dots, P_i + \Delta P_i, \dots P_N) - F(P_1, P_2, \dots, P_i, \dots P_N)}{\Delta P_i} \quad (15)$$

Where N is the number of parameters considered, F is the dependent variable and P is the independent parameter. A negative coefficient indicates an inversely proportional relation between a dependent variable and an independent parameter. By normalizing the sensitivity coefficient using reference values representing the ranges of the dependent variable and independent parameter the sensitivity index can be calculated. It is given as (Gu *et al*, 2002):

$$S_i = \frac{P_m}{F_m}\frac{\Delta F}{\Delta P} \tag{16}$$

Where F_m and P_m are the mean of the highest and the lowest values of the range for respectively the explanatory variable and the explanatory parameter. A negative sign shows inverse proportionality and the sensitivity index indicates a higher sensitivity when it has a higher absolute value.

Based on extensive literature review (Ahmad, Jamieson, Havard, Madani, & Zaman, 2009; Feyereisen, Strickland, Bosch, & Sullivan, 2007; Merz & Blöschl, 2004; Spruill, Workman, & Taraba, 2000; White, Harmel, Arnold, & Williams, 2014) and studying the theoretical documentation of the SWAT model (Neitsch et al., 2009), eight input parameters were selected for sensitivity analysis which strongly determine discharge; the curve number (CN), plant uptake compensation factor (EPCO), soil evaporation compensation factor (ESCO), soil available water capacity (SOL_AWC), baseflow alpha factor (ALPHA_BF), groundwater revap (GW_REVAP), deep aquifer percolation coefficient (RECHRG_DP) and groundwater delay time (GW_DELAY). Table 9 shows the ranges and initial estimates for the selected parameters. More information about the model parameters can be found in paragraph 2.6 or in Arnold *et al (*2012) and Neitsch *et al* (2009). The model determines the initial values for each parameter based on the average and most suitable value. The SWAT model parameters are mostly process-based. Runoff curve number and the evaporation coefficients are two exceptions. Therefore these parameter values must be within their applicability ranges.

Table 9. Parameter ranges and initial values used in the sensitivity analysis

Model parameter	**Variable name**	**Range**	**Initial model estimates**
Curve Number	CN	65-88	77
Soil evaporation compensation factor	ESCO	0.7-0.95	0.95
Plant evaporative compensation factor	EPCO	0.1-1	1
Soil available water capacity (mm)	SOL_AWC	±0.05	-
Baseflow alpha factor	ALPHA_BF	0.05-0.8	0.048
Groundwater revap coefficient	GW_REVAP	0.02-0.2	0.02
Groundwater delay time (days)	GW_DELAY	0-150	31
Deep aquifer percolation fraction	RCHRG_DP	0-1	0.05

The sensitivity analysis uses the model parameters as independent or explanatory variable and baseflow and quick flow as the dependent variables. The sensitivity indices and coefficients were examined for both the local scale model and the regional scale model to characterize quick flow and baseflow under different parameter ranges. The SWAT model separates three different types of flow that can enter into a stream. Surface runoff, represents overland flow. Lateral flow, represents flow through the soil layer. Return flow, represents flow from the unconfined aquifer and deep aquifer into the stream. For the sensitivity analysis lateral flow and surface runoff are both considered as the fast component and return flow is considered as the slow component. As the model calibration will be done using monthly discharge values lateral flow and surface runoff are summed and defined as quick flow. Return flow is defined as baseflow.

Table 10 shows the sensitivity coefficients and sensitivity indices for the parameters that correspond to change in quickflow and baseflow volumes resulting from changing the model parameters in the local watershed model and the regional watershed model. Sensitivity analysis gives insight in the dominant processes determining model output and therefore critical for successful calibration and application of models. Sensitivity of a parameter in a watershed can be significantly different in another watershed. Scale differences between watersheds can be of importance. Therefore it is imperative that results from sensitivity analysis are examined before calibration starts.

Table 10. Sensitivity indices of SWAT parameters for the local watershed and the regional watershed model. Indicated in the left are the independent variables, i.e. the parameters. The dependent variables are baseflow and quickflow. The values in the

bold format indicate the sensitivity index value of baseflow and quickflow for each parameter. Negative values for S_i indicate inverse proportionality and higher absolute values indicate higher sensitivity of the dependent variable to parameter changes.

Local watershed model																	
Parameter	Initial value	Independent variable				Dependent variable (Baseflow)						Dependent variable (Quickflow)					
		P1	P2	ΔP	Pm	F1	F2	ΔF	Fm	ΔF/ΔP	(Pm/Fm)*(ΔF/ΔP) = Si	F1	F2	ΔF	Fm	ΔF/ΔP	(Pm/Fm)*(ΔF/ΔP) = Si
SOL_AWC	-	0.04	0.04	0.08	0.04	175.1	248.9	-74	.12	-921.7	**-0.17**	426	477	-52	452	-645.2	**-0.06**
CN	77	85	69	20	77	38.7	333.7	-295	.86	-14.7	**-6.10**	567	319	248	443	12.4	**2.16**
GW_REVAP	0.02	0.02	0.2	0.18	0.11	202.8	175.1	28	.89	153.6	**0.09**	451	451	0	451	0.0	**0.00**
GW_DELAY	31	0	100	100	50	199.1	195.4	4	.97	0.0	**0.01**	459	459	0	459	0.0	**0.00**
ESCO	0.95	0.5	1	0.5	0.75	127.2	202.8	-76	.65	-151.2	**-0.69**	394	462	-67	428	-134.6	**-0.24**
RCHRG_DP	0.05	0	1	1	0.5	208.3	167.7	41	.88	40.6	**0.11**	459	459	0	459	0.0	**0.00**
EPCO	1	0.01	1	0.99	0.505	228.6	202.8	26	.16	26.1	**0.06**	489	459	30	474	30.7	**0.03**
ALPHA_BF	0.048	0.048	0.8	0.75	0.424	202.8	210.1	-7	.06	-9.8	**-0.02**	475	459	16	467	21.3	**0.02**
Regional watershed model																	
Parameter	Initial value	Independent variable				Dependent variable (Baseflow)						Dependent variable variable (Quickflow)					
		P1	P2	ΔP	Pm	F1	F2	ΔF	Fm	ΔF/ΔP	(Pm/Fm)*(ΔF/ΔP) = Si	F1	F2	ΔF	Fm	ΔF/ΔP	(Pm/Fm)*(ΔF/ΔP) = Si
SOL_AWC	-	0.04	0.04	0.08	0.04	165.9	262.7	-97	.14	-1209.7	**-0.23**	428	477	-50	453	-622.1	**-0.05**
CN	77	85	69	20	77	38.7	387.1	-348	.13	-17.4	**-6.30**	581	304	277	442	13.8	**2.41**
GW_REVAP	0.02	0.02	0.2	0.18	0.11	202.8	175.1	28	.89	153.6	**0.09**	459	459	0	459	0.0	**0.00**
GW_DELAY	31	0	100	100	50	201.8	195.4	6	.99	0.1	**0.02**	459	459	0	459	0.0	**0.00**
ESCO	0.95	0.5	1	0.5	0.75	127.2	210.1	-83	.69	-165.9	**-0.74**	387	470	-83	429	-165.9	**-0.29**
RCHRG_DP	0.05	0	1	1	0.5	221.2	167.7	53	.94	53.5	**0.14**	459	459	0	459	0.0	**0.00**
EPCO	1	0.01	1	0.99	0.505	235.0	202.8	32	.19	32.6	**0.08**	487	442	44	465	44.7	**0.05**
ALPHA_BF	0.048	0.048	0.8	0.75	0.424	207.4	210.1	-3	.09	-3.7	**-0.01**	480	442	38	461	50.1	**0.05**

Both models show the same tendencies looking look at the results of the sensitivity analysis [S_i = *(Pm/Fm)*(ΔF/ΔP)*]. Four parameters show the largest model sensitivity. From largest to smallest; CN, ESCO, RCHRG_DP and SOL_AWC. Of which CN and ESCO are inversely proportional for baseflow. RCHRG_DP shows positive sensitivity for baseflow and no sensitivity quickflow. SOL_AWC is negatively proportional for both baseflow and quickflow. Other parameters have a value of lower than the absolute Si value of 0.1 and are considered as not significantly sensitive to adjustments.

9.2 Graphical and statistical analysis

After applying the influence coefficient method the SWAT Check program is used to attain more insight in the influence of parameters on the model output. SWAT Check generates a graphic showing water flow per component of the water balance. The value is the yearly sum of water in mm averaged over the period the SWAT model is run. Figure 37 shows an example of SWAT Check's output. The output graphic shows various reservoirs defined within the model, such as the soil layer, shallow aquifer and deep aquifer.

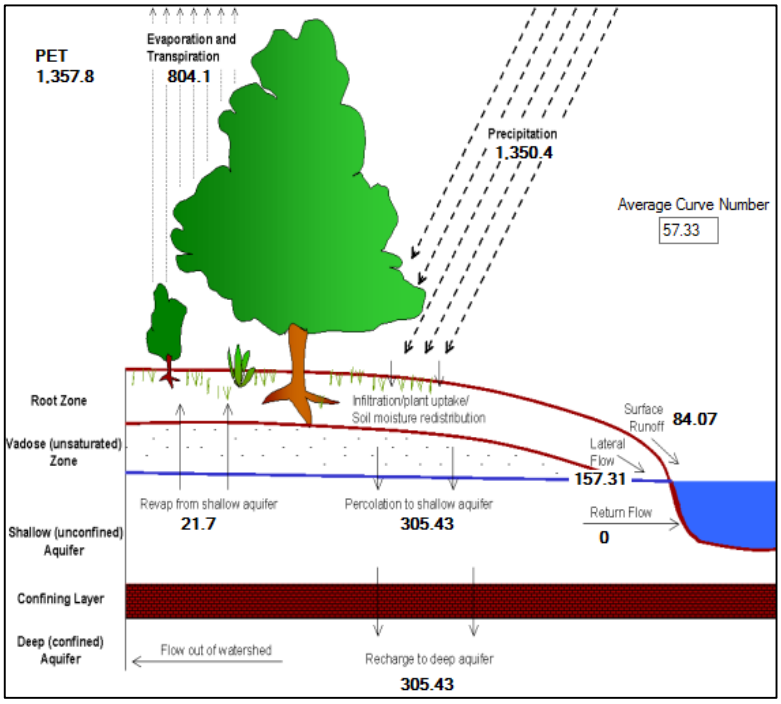


Figure 37. Screenshot of SWAT Check output example (White et al., 2014)

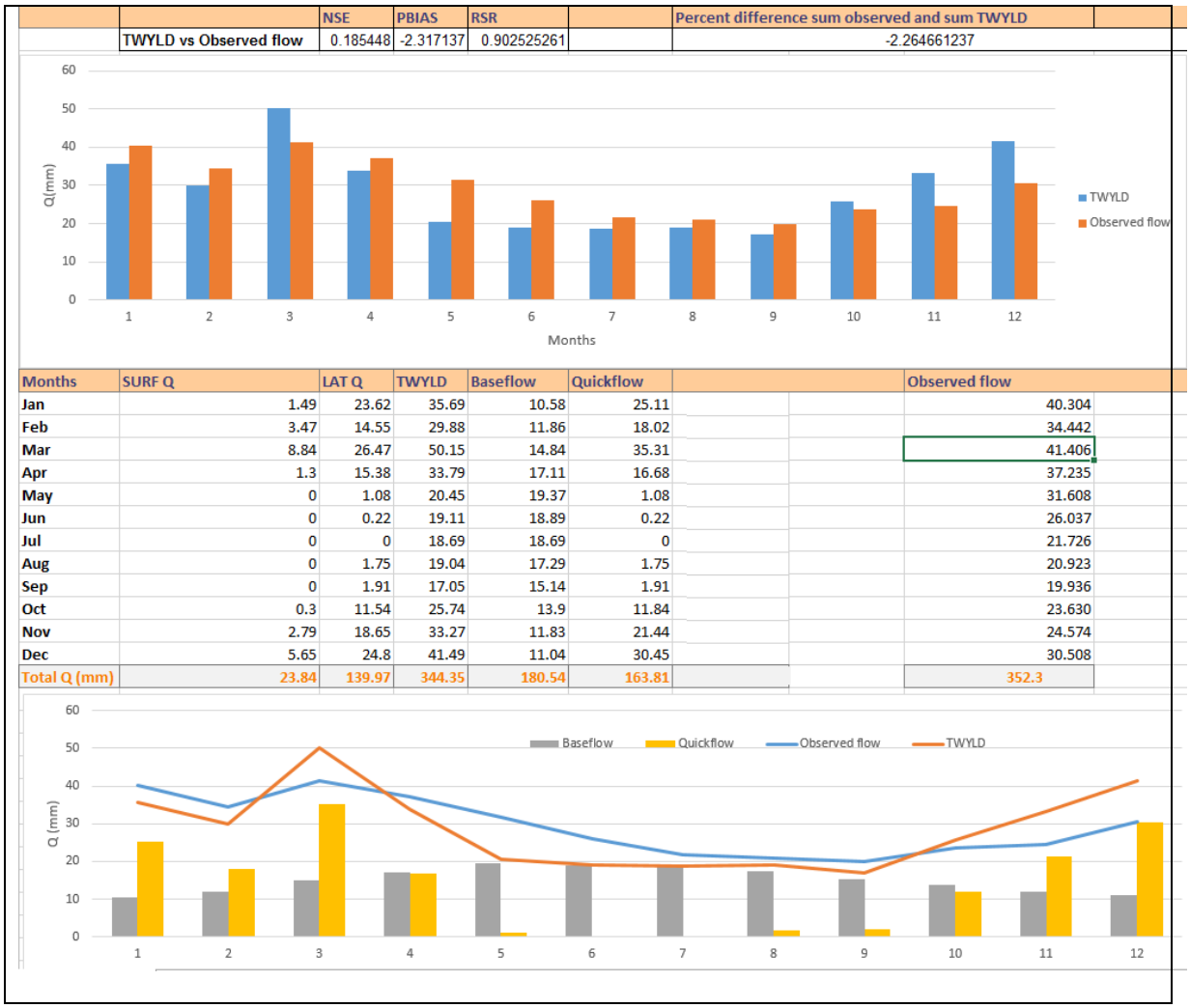
Moreover flow directions are shown, e.g. groundwater revap flows from the shallow aquifer into the root zone due to capillary rise. The program is developed by White, *et al* (2014) and functions as a screening tool that helps users to identify potential problems in the SWAT model. The program also shows values for precipitation, snow fall, surface runoff, lateral flow, water yield, evapotranspiration, sediment yield and potential evapotranspiration, shown in table 11. Water yield is the total amount of water that exits the watershed. In the previous paragraph is explained that lateral flow and surface runoff are both considered as quickflow. Baseflow is thus the difference between total water yield (TWYLD) and quickflow. The first column of table 11 represents the months in a year, the neighboring columns show average values of the sum of water in mm for the entire period the SWAT model is run. The table shows an example of the regional model. It is a good sign that in this tropical climate during the entire period snow fall is 0.0 mm.

Table 11. SWAT Check monthly average sum of components of the water balance (White et al., 2014)

Mon	Rain (MM)	Snow Fall (MM)	SURF Q (MM)	LAT Q (MM)	Water Yield (MM)	ET (MM)	Sed. Yield (MM)	PET (MM)
1	189.66	0.00	12.71	23.48	54.08	77.01	0.14	117.68
2	221.42	0.00	21.10	29.08	75.93	70.25	0.26	97.53
3	186.99	0.00	18.77	24.97	84.67	85.18	0.21	109.26
4	95.82	0.00	7.89	9.90	63.84	83.17	0.08	105.74
5	22.67	0.00	0.23	2.02	46.76	78.00	0.00	106.15
6	8.26	0.00	0.04	0.65	35.88	55.42	0.00	97.69
7	9.71	0.00	0.40	0.80	29.39	45.12	0.00	116.01
8	17.01	0.00	0.00	1.22	22.42	46.16	0.00	135.37
9	30.69	0.00	0.05	2.55	17.90	51.74	0.00	140.51
10	94.28	0.00	0.58	8.48	20.78	65.76	0.00	128.96
11	221.24	0.00	5.72	23.20	37.35	71.57	0.03	105.75
12	250.85	0.00	16.40	30.74	56.54	74.19	0.11	96.41

SWAT Check is initially created to check mistakes in the SWAT model, but in this research the SWAT Check output is used for calibration as well. In order to do this a tool was developed for this research using excel and Visual Basic. The tool helps to graphically and statistically interpret the SWAT Check output. The program generates two graphs showing monthly baseflow, quickflow and TWYLD. As well as results for the performance indicators, e.g. Nash Suthcliffe Efficiency (NSE), percent bias (PBIAS), Ratio Root mean square error and standard deviation (RSR). The tool is named SWAT Check Discharge Interpreter (SWAT-CDI) and requires the user to copy and paste the fourth, fifth and sixth row from the SWAT Check output

table into the first column and insert monthly sum values for the observed discharge in the seventh column. The excel tool then automatically generates the FIGUREs and calculates the performance indicator values. Calculating the average monthly sum discharge values from the calibration or validation time series can easily be done using the pivot tool functions of Excel.



	NSE	PBIAS	RSR	Percent difference sum observed and sum TWYLD
TWYLD vs Observed flow	0.185448	-2.317137	0.902525261	-2.264661237

Months	SURF Q	LAT Q	TWYLD	Baseflow	Quickflow	Observed flow
Jan	1.49	23.62	35.69	10.58	25.11	40.304
Feb	3.47	14.55	29.88	11.86	18.02	34.442
Mar	8.84	26.47	50.15	14.84	35.31	41.406
Apr	1.3	15.38	33.79	17.11	16.68	37.235
May	0	1.08	20.45	19.37	1.08	31.608
Jun	0	0.22	19.11	18.89	0.22	26.037
Jul	0	0	18.69	18.69	0	21.726
Aug	0	1.75	19.04	17.29	1.75	20.923
Sep	0	1.91	17.05	15.14	1.91	19.936
Oct	0.3	11.54	25.74	13.9	11.84	23.630
Nov	2.79	18.65	33.27	11.83	21.44	24.574
Dec	5.65	24.8	41.49	11.04	30.45	30.508
Total Q (mm)	23.84	139.97	344.35	180.54	163.81	352.3

Figure 38. SWAT Check Discharge Interpreter output. The upper rows show the results of three performance indicators; Nash-Suthcliffe Efficiency, Percent Bias and Ratio Root mean square error and Standard deviation. The percentage difference in the sum of observed and simulated discharge is shown in the upper right cell. The upper column graph shows the difference in TWYLD (total water yield, i.e. discharge) and the observed flow (observed monthly average discharge) in mm. The table in the middle shows the monthly summed values for surface flow (SURF Q), lateral flow (LAT_Q), total water yield (TWYLD), baseflow, quickflow and observed flow. Underneath the monthly time series the sum of the whole year of each output parameter is shown. In the lower graph monthly average baseflow, quickflow, observed flow and TWYLD are figured.

The upper graph in figure 38 shows how monthly summed TWYLD corresponds to the monthly summed observed discharge. The lower graph shows the components quick flow and baseflow to the TWYLD. And above the upper graph NSE, PBIAS and RSR values are shown.

Using the program SWAT Editor, which contains the tool 'Manual Calibration Helper' it is possible to easily change parameter values in the SWAT model. By manually changing SWAT parameters within realistic ranges using SWAT editor, then analyzing flow component changes in SWAT Check (figure 37 and table 11), analyzing graphical and statistical results from SWAT-CDI (figure 38) and lastly doing extensive literature research of the theoretical documentation of the SWAT model to gain deeper insights into the workings of the model could be attained. These insights are described in table 13. It shows how the model reacts and why when parameters increase. The parameters Saturated Hydraulic Conductivity (SOL_K) and Soil Layer Depth (SOL_Z) were also varied during this analysis.

TABLE 12. Parameter ranges of soil hydraulic conductivity (SOL_K) and soil layer depth (SOL_Z)

Model Parameter	Variabel name	Range
Saturated hydraulic conductivity (mm/hr)	SOL_K	+- 40%
Soil layer depth from surface (mm)	SOL_Z	+- 20%

Table 12 shows the realistic ranges in between which sol_k and sol_z were varied. The reason these parameters were also assessed was because there was a high spatial variability of these values in the regional and local watershed.

Table 13. Descriptions of model response to an increase of the model parameters. Left column shows the abbreviation of the parameter, the middle column the name and if applicable unit and the right column explains how either discharge, baseflow and quickflow respond to model parameters increases.

Parameter	Variable name	Model response to increasing the model parameter
SOL_AWC	Soil available water content (mm H2O/mm soil)	More water is available in the soil, which means more water can be absorbed by plant roots, resulting in a higher evapotranspiration. Less water enters the stream via lateral flow, thus quick flow decreases. Little influence for baseflow
CN	Curve Number	Surface runoff increases, resulting in less water infiltration. This reduces lateral flow and percolation from the soil zone to the unsaturated zone, thus decreasing baseflow.
GW_REVAP	Groundwater Revap	This results in more water moving upward from the shallow aquifer to the soil layer, which increases lateral flow and increases water availability in the root zone, resulting in higher root water uptake and thus more water for transpiration.
GW_DELAY	Groundwater Delay (days)	This does not influence the amount of baseflow, but it increases the time it takes for water to reach the stream. Higher GW_DELAY means that more water becomes available in the dry season in the form of baseflow.
ESCO	Soil Evaporation Compensation Factor	This increases evapotranspiration, resulting in a reduction of the total water yield in the river. It both reduces surface runoff as baseflow.
EPCO	Plant Evaporation Compensation factor	Changing this parameter had very little effect on the two models. Increasing EPCO to the highest value in the range, slightly increased baseflow and surface runoff.
RECHARGE_DP	Deep Aquifer percolation factor	Increasing RECHARGE_DP did not change the amount of baseflow. Increasing this parameter means that more water moves from the shallow aquifer into the deep confined aquifer. Water moving through this aquifer reaches the river the slowest. Increasing this parameter thus resulted in more water entering the stream in the dry period.
ALPHA_BF	Baseflow alpha factor	Increasing this parameter means more water enters the stream via the shallow (unconfined) aquifer and less water percolates into the deep confined aquifer. Increasing this parameter, thus reduces the amount of water entering the stream in the dry period.
SOL_K	Saturated hydraulic conductivity (mm/hr)	Higher SOL_K means water moves quicker through the soil layer and into the stream, thus increasing lateral flow and less water can percolate into the shallow aquifer and the deep aquifer, resulting in less baseflow.
SOL_Z	Soil layer depth from surface (mm)	Increasing the soil layer depth increases the amount of water available in the root zone and thus increases evapotranspiration, reducing total water yield. A thicker soil layer also increases lateral flow and thus reduces baseflow.

9.3 Calibration theory

There are various problems that occur during calibration of distributed watershed models. An important question to ask yourself is 'when is a model calibrated?'. When a watershed outlet is shows correct discharge values, does this also mean that discharge in the tributaries inside the watershed show correct values? It is also important to define the goal and purpose of the model. The SWAT model is capable of assessing water quality, sediment transport, discharge and other components of the hydrological cycle. The parameter choice for calibration is dependent on the output of interest, which depends on the goal of the model. In this research we are only interested in monthly average discharge values. A second problem is calibrating a watershed that is highly managed. Perhaps natural processes play a secondary role in these watershed. If detailed management data, such as irrigation data, extraction wells or hydro-electric dams, is not available, then how well is the model capable of simulating discharge? Uncertainty in the model is a third problem. Model uncertainty can be divided into uncertainty about the conceptual model, input uncertainty and parameter uncertainty. Oversimplification in conceptual models can ignore important processes. Input data can contain mistakes or missing data. And parameters, which represent processes, can neutralize each other's influence to produce non-unique parameter set. Objective functions used during calibration can give good results for different parameter sets, called multimodality, otherwise called the Swiss cheese effect (Abbaspour, 2016). All these problems need to be kept in mind during calibration and it is important to realize that the final calibration adjustments will not be unique. The main goal of calibration in this research is to attain a model that is able to produce monthly average discharge values at the watershed outlet.

An often used calibration methodology in SWAT papers is the method described by Neitsch *et al*, (2002), which is also referred to in the SWAT manual. This method calibrates the model first on average annual discharge, then monthly average discharge and finally daily average discharge. Moreover it advises to separate observed discharge into quickflow and baseflow. The logic behind this results from the chronological order in which water moves through the hydrological cycle and eventually leaves the watershed as discharge. After a rainfall event water will first reach the river as surface runoff, remaining water will infiltrate and reach the river as lateral flow and as part of the water percolates to deeper layers eventually water will reach the river as baseflow. Quickflow defined as the sum of surface runoff and lateral flow reaches the river earlier then baseflow and thus values for simulated quickflow need to be calibrated first according to this methodology Neitsch *et al*, (2002). There are various baseflow separation methods, of which the baseflow filter developed by Arnold *et al* (1995) is the most used filter in SWAT papers.

To assess goodness-of-fit during calibration there are a multitude of quantitative statistics that can be applied. Moriasi *et al.*, (2007) analysed statistical evaluation techniques for simulating discharge and recommend using Nash-Suthcliffe efficiency (NSE), percent bias (PBIAS) and ratio of the root mean square error (RSR) in combination with graphical techniques (e.g. graphs, boxplots) for model calibration. NSE results in a value below 1. The closer NSE is to 1, the better the fit. NSE is a correlation based measure, which is oversensitive to extreme values (outliers) and is insensitive to proportional differences between model prediction and observations (Legates & McCabe, 1999). Because of this limitation of NSE it is good to use extra performance indicator. PBIAS is not a correlation based measure and does take proportional differences between observed and simulated values into account. PBIAS gives percentage differences, 0% being a perfect fit. RSR is an alternative to Root mean square error (RMSE), a very commonly used error index statistic (Singh, Bankar, *et al.*, 2013). It calculates the ratio of RMSE and standard deviation of measured data. The value for RSR becomes perfect as the value approaches 0. Moriasi *et al* (2007) state that discharge simulations can be judged satisfactory if NSE>0.50 and RSR $\leq$ 0.70, and if PBIAS $\pm$25%, equations 17, 18 and 19 show their equations respectively. Q_m is defined as measured discharge and Q_s is simulated discharge.

$$NSE = 1 - \frac{\sum_i (Q_m - Q_s)_i^2}{\sum_i (Q_{m,i} - \overline{Q}_m)^2} \tag{17}$$

$$PBIAS = 100 * \frac{\sum_{i=1}^{n} (Q_m - Q_s)_i}{\sum_{i=1}^{n} Q_{m,i}} \quad (18)$$

$$RSR = \frac{\sqrt{\sum_{i=1}^{n} (Q_m - Q_s)_i^2}}{\sqrt{\sum_{i=1}^{n} (Q_{m,i} - \overline{Q}_m)^2}} \quad (19)$$

Besides manual calibration methods, automatic calibration methods are also popular, e.g. GLUE, ParaSOL, MCM, PSO or SUFI2. In SWAT papers recently the program SWAT-CUP is used, which is developed by Abbaspour *et al* (2015). It is a relatively new program that allows users to automatically calibrate using different methodologies. The most popular one is SUFI2, which stands for Sequential Uncertainty Fitting. Uncertainty in parameters in this method is expressed as ranges. The uncertainty is expressed as the 95% probability distributons. Latin hypercube sampling is used to generate the propagation of parameter uncertainties, which is used to calculate the 95% probability distribution of the output variable (95PPU). The 95PPU expresses an envelope of good solutions generated by certain parameter ranges (Abbaspour, 2015). To show the fit between simulations and observations two statistics are used, the P-factor and the R-factor. The P-factor is the percentage of observed values in the 95PPU envelope and the R-factor is the thickness of the 95PPU envelope (Abbaspour, 2015). A good fit is accepted if the P-factor > 0.7% and the R-factor around 1. SUFI2 performs several iterations, which in SWAT-CUP can be previously defined. After each iteration SUFI2 zooms in on a region of parameter space, which produces better results than the previous iteratirion, resulting in a larger P-factor and a smaller R-factor. A P-factor of 1 and an R-factor of zero is a simulation that exactly corresponds to measured data. The degree in which values are away from these numbers can be used to judge the calibration (Abbaspour, 2015).

In this research the choice was made to first manually calibrate the model according to the Neitsch method, adjusting the parameters which showed the most sensitivity described in the previous paragraph. Because in this research we are interested in the long term assessment of climate change and land-use change a calibration on a daily time step resolution will not be done. After the models show an acceptable result after manual calibration. Automatic calibration is performed using the SWAT-CUP SUFI2 method. Abbaspour (2015) advises on using at least 500 iterations for the automatic calibration. As input for SWAT-CUP for each of the sensitive parameters, the manually calibrated parameter serves as the initial value and a realistic range is selected per parameter.

9.4 Baseflow separation

There are various hydrograph separation techniques to identify various flow components from total flow. The components represent various flowpaths, with each their own resiedence time. Quickflow has a shorter residence time than baseflow. There are numerical analytical separation methods that use computer programs for separation. Additionally there are isotope and chemical hydrograph separation techniques that use tracers to determine the separation. The recession curve of a hydrograph can be used to determine baseflow separation. And lastly dynamic groundwater table depth measurements in combination with head water values can give an indication of baseflow separation.

Numeric analytical methods separate the hydrograph in a fast and a slow component. For SWAT researches the developer of SWAT Arnold *et al* (1995) developed a program that does this automatically, called the Baseflow Program. This method uses the digital filter method of Nathan & McMahon (1990) and is based on the equation:

$$q_t = \beta q_{t-1} + \frac{1-\beta}{2} * (Q_t - Q_{t-1}) \quad (20)$$

where β is the filter parameter (0.925, q_t is surface runoff at time step t (one day), Q is the original discharge. The value of β was determined by Nathan and McMahon (1990) to give realistic results.

$$b_t = Q_t - q_t \tag{21}$$

The model can pass the time series three times to determine baseflow, forward, backward and forward again. Every pass results in more quickflow and less baseflow. The pass used for the determination of the baseflow component should be motivated by other criteria, such as the other separation methods described above. An important criteria for this research is the depth of the waterTABLE.

The Institute of Hydrology introduced the Base Flow Index (BFI) in 1980, which calculates the ratio of baseflow from total flow after smoothing and separation procedures using daily discharges. Values for BFI range from 0 to 1. Higher values represent catchments that have higher infiltration and thus higher base flow. Morawietz (1997) developed a BFI program, using Visual Basic in Excel, which allows users to insert daily discharge values to attain baseflow separation from the hydrograph.

Figure 39 shows a comparison between different passes of the Base Flow Program of Arnold and the BFI filter program of Morawietz, using daily average discharge measurements from the Buriti Vermelho river. For calibraton the choice was made to use the BFI filter program, because it produced results for separation that were in between the results from the Base Flow Program. Research, using other methods, such as isotopes or recession analysis, was not found in literature for the Preto river or Buriti Vermelho river.

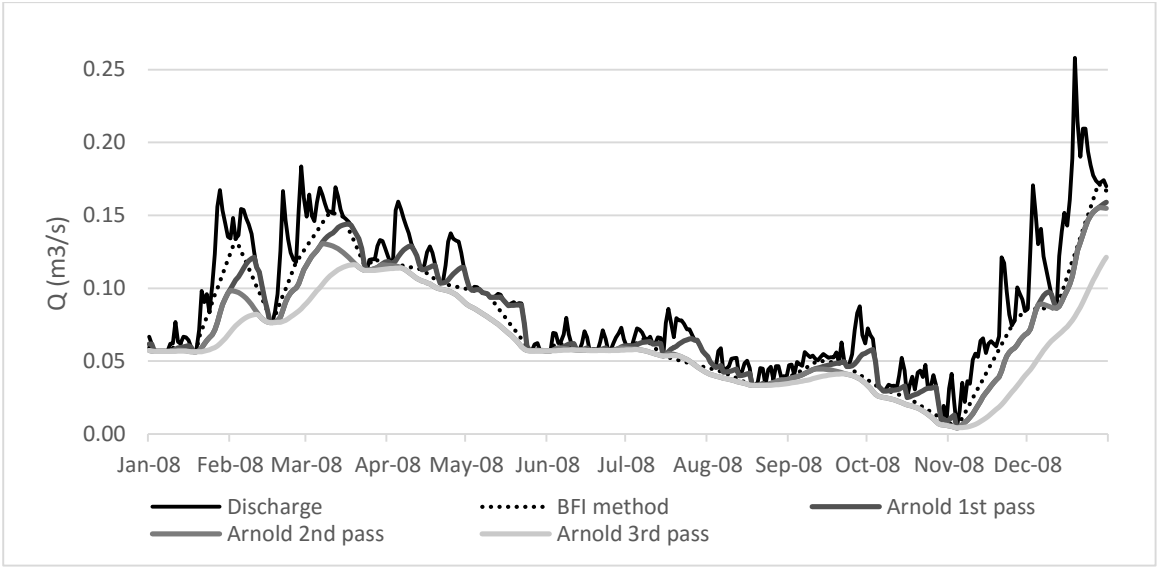


Figure 39. Daily average measured discharge at the Buriti Vermelho watershed outlet. Three passes (forward, backward and forward) with the Base Flow program (Arnold et al., 1995) *and one pass using the BFI filter* (Morawietz, 1997).

10. Calibration

10.1 Calibration and validation period

The first step in calibrating a watershed model is dividing the discharge time series in a calibration and validation period. In the calibration period parameter values can be varied until an accepTABLE model simulation fit is obtained. Then the model can be run with the same parameter adjustments for the validation period. Finally a goodness-of-fit is determined, using an appropriate performance indicator. It is important to choose a calibration and validation period, which both contain a wet and a dry period, because this shows that the model is able to account for this natural variability in precipitation.

The regional watershed model uses discharge data from the Fazenda Limeira hydrological station in a period where there was little irrigated agriculture and where the Queimados dam was not built yet. This results in discharge time series that represent the natural system of the regional watershed. The chosen calibration period is between January 1986 and October 1991, containing the dry years of 1987 and 1989, the average years of 1988 and 1991 and the very wet year of 1990 (figure 23). The years before 1986 are not selected

because of a large missing period between July 1985 until December 1985 and high percentage of missing data between January 1984 until July 1986 of >10%.

The local watershed model will not be calibrated using its own observed discharge data set, because the calibration will be done using the downscaling methodology. The discharge time series thus does not have to be split up. The entire discharge time series can be used for validation of the downscaling methodology.

10.2 Calibration regional watershed model

The first run of the regional watershed model produced the graph shown in figure 40, which shows monthly average simulated discharge and observed monthly average discharge for the calibration period.

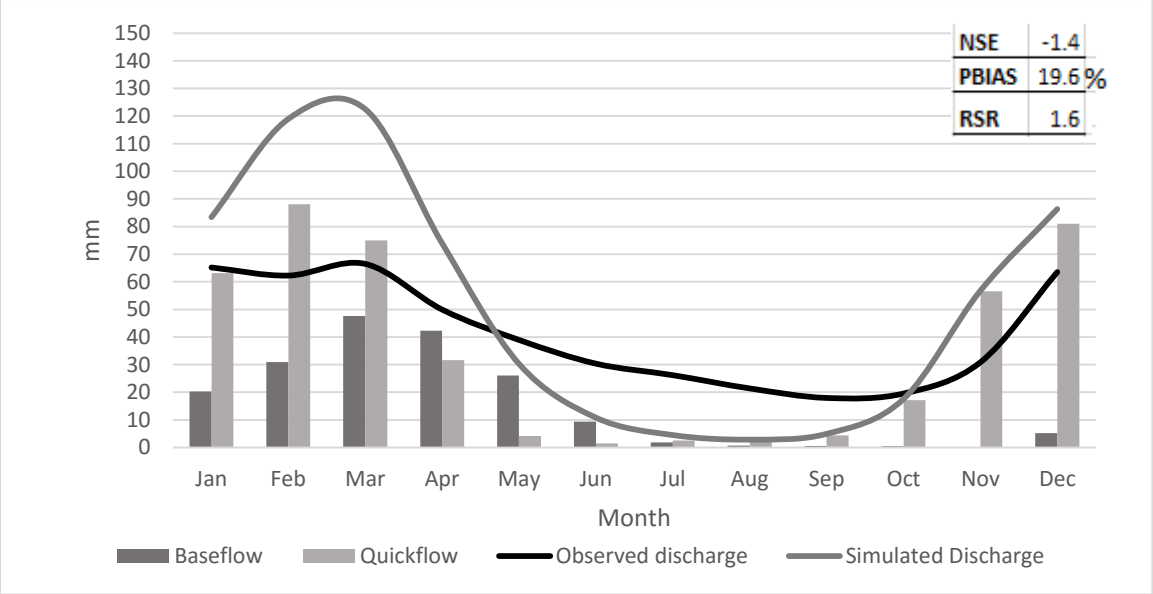


Figure 40. First result of the regional watershed model using the initial values determined by SWAT. The lines show average simulated monthly discharge and observed discharge. The columns represent baseflow and quickflow as simulated by SWAT. In the upper right corner the performance indicators NSE, PBIAS, PERR and RSR indicate the goodness-of-fit of the average simulated discharge.

Three important observations can be made from figure 41. The NSE value is negative, i.e. the model simulates monthly values worst than the average of the observed time series would. The PBIAS is 19.6%, i.e. average monthly discharge values are ±20% too high. The total sum of yearly discharge must thus be lowered. Lastly there is too much discharge in the wet period and too little discharge in the dry period. Figure 41 shows the monthly average discharge values over the entire calibration period. A positive observation is that he peaks are generally at the same places. The whole period has an NSE of -0.48 and a PBIAS of 22.3%. The table in the left upper corner of figure 41 shows the NSE and PBIAS per year. The year 1989 has a positive NSE value. This probably results from this year being a wet year. In this initial stage the model seems to be able to predict wet years better.

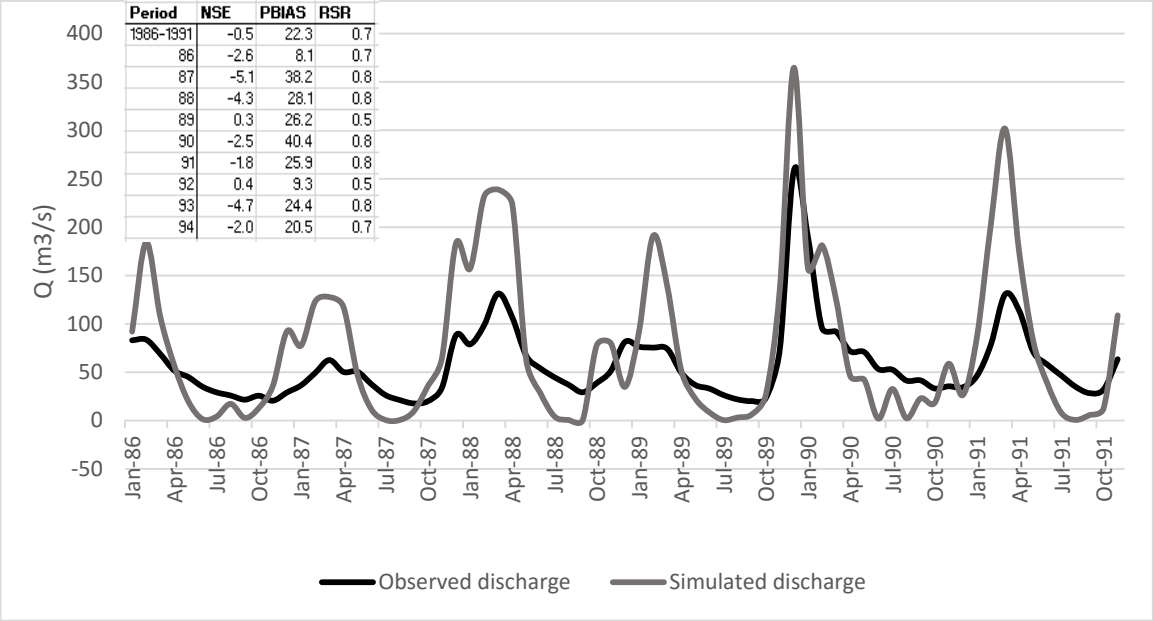


Figure 41. First result of the regional watershed model using the initial values determined by SWAT. The lines show average simulated monthly discharge and observed discharge. In the upper left corner the performance indicators NSE, PBIAS, and RSR indicate the goodness-of-fit of the average simulated discharge for the whole calibration period (first row) and for each individual year (2nd row onwards).

In order to carry out the calibration according to the Neitsch *et al*. (2002) method the BFI filter program is used to separate baseflow and quickflow from the total discharge measured at the Fazenda Limeira hydrological station. The BFI filter requires daily discharge values and the output is also in daily values. Using the watershed area and a pivot table in Excel monthly summed values are calculated mm. The resulted graph is shown in figure 42.

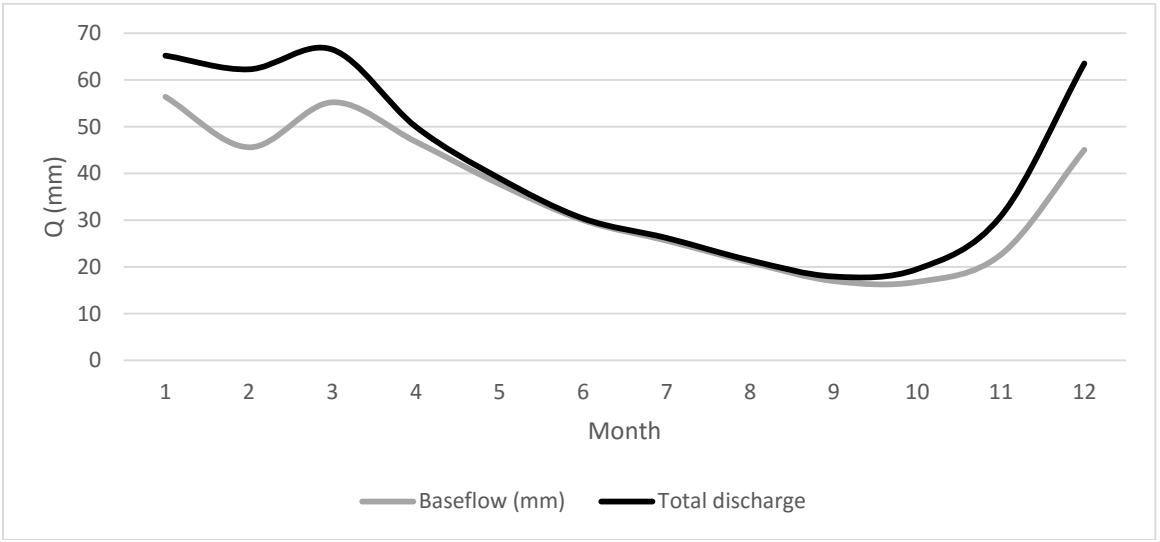


Figure 42. Monthly average discharge and baseflow per month in mm calculated for the calibration period 1986 untill 1991 using observed discharge data from the Fazenda Limeira hydrological station. Baseflow is generated using the BFI filter developed by (Morawietz, 1997)

The lower graph in the SWAT-CDI excel tool was used during the calibration (figure 38). The first step of the Neitsch (2002) method is to calibrate yearly average simulated discharge values to fit with the yearly average observed discharge values. Simulated average discharge values are too high, as can be seen from the PBIAS values in figure 41. These values can be lowered by adjusting the ESCO parameter, which changes the amount of water available for soil water evaporation. The next step according to Neitsch et al. (2002) is to attain quickflow values which are similar to quickflow values attained with the BFI filter. The ratio of quickflow to baseflow in the initial model results (figure 40) is much higher than the ratio determined by the BFI filter (figure 42). In order to reduce simulated quickflow the paramters SOL_K, SOL_Z and CN2 can be

reduced. If simulated quickflow is reduced to the point where a baseflow/quickflow ratio is attained which is similar to the ratio produced by the BFI filter, the values for SOL_K, SOL_Z and CN2 have to be reduced to an unrealistically low value. The baseflow/quickflow ratio from the baseflow filters thus does not represent the hydrology of the regional watershed properly.

At this point in the calibration process the validity of the baseflow filter programs is questioned and their usability in a tropical climate with the typical Cerrados soil. Four arguments against the use of the baseflow filters in this typical environment are presented. Firstly, the SWAT model, which to a certain extent is able to reproduce reproduce realistic processes is unable to produce the results of the baseflow filter, without using unrealistic parameter values. Secondly, the unique tropical climate, with a very wet period and a completely dry period with no precipitation. Thirdly, the Cerrados soil, which is characterized by a very deep soil layer (± 6 meter) which has a very high saturated hydraulic conductivity. This suggest that a large portion of the water in the wet season percolates to deeper aquifers. And lastly the deep groundwater level, which varies between 6 and 20 meters deep.

Delayed baseflow response in tropical Cerrados watersheds

Based on the previous arguments a theory about the behavior of water in this tropical Cerrados region is formulated. During the wet season the deep aquifer is recharged resulting from a surplus in precipitation. A large part of discharge results from quickflow. During these wet months the groundwater table rises slowly to a point that the Preto river drains more water from this groundwater storage. When the wet season ends the groundwater table is at its highest point. No further recharge is expected, because travel times are fast in the soil layer because of the high saturated hydraulic conductivity. From this point river discharge only consists of water drained from the groundwater storage. Because of the high temperature and solar radiation the soil layer tends to dry out, resulting in a gradient from the groundwater to the soil layer. This results in capillary rise, which increases the groundwater table. As the dry period months go by, the groundwater storage reduces to the point where the first wet months start again. The projected baseflow and quickflow based on this theory is presented in figure 43 in the left graph. The right graph of figure 43 shows the results of the baseflow filter. The main difference is that quickflow is a larger component in the wet months than baseflow. Quickflow is the summation of surface runoff and lateral flow.

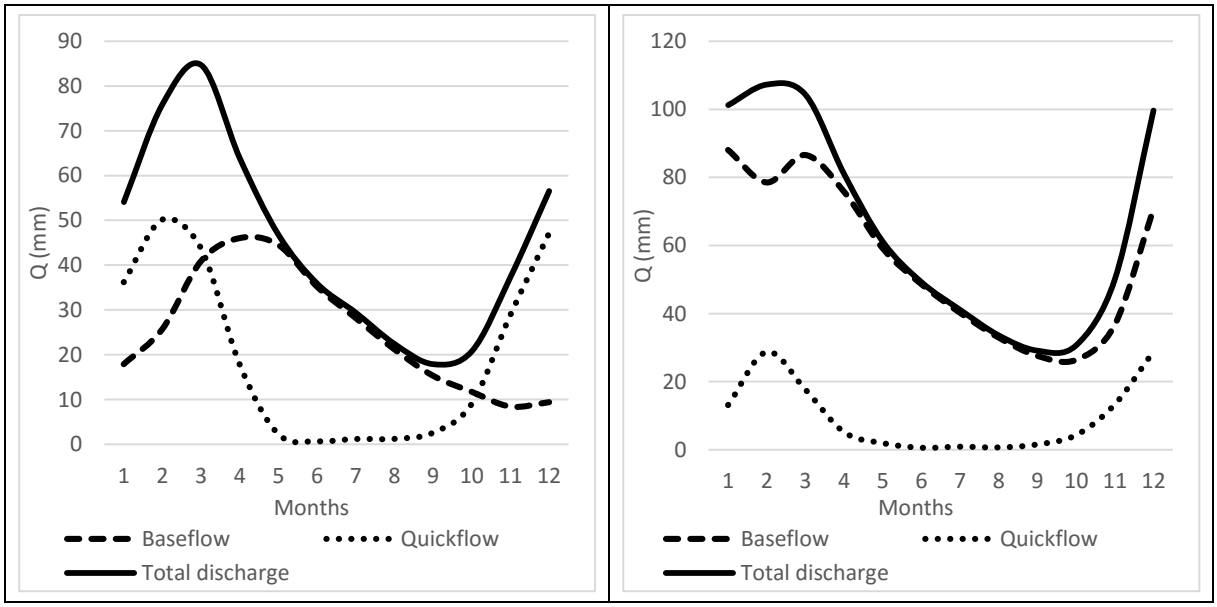
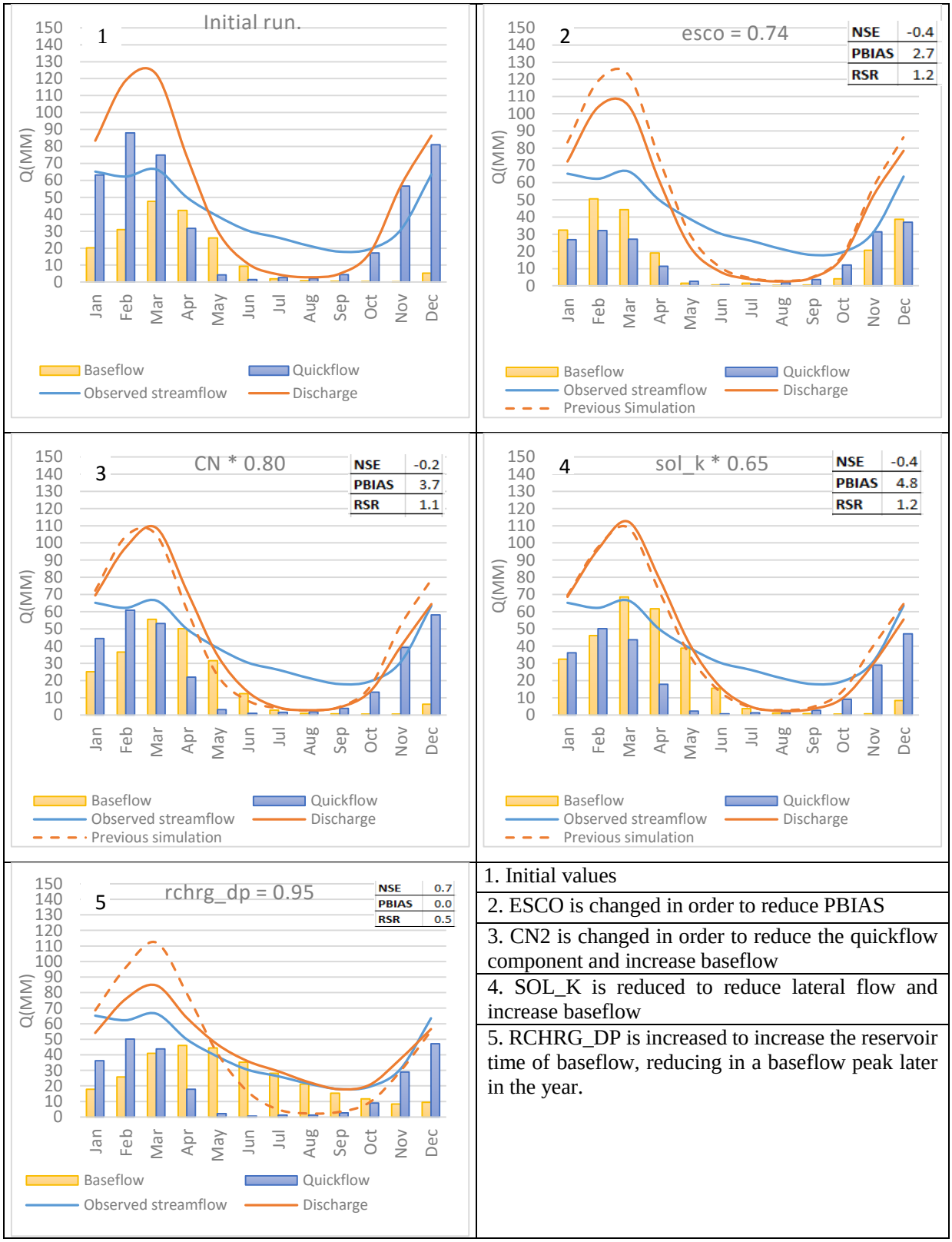


Figure 43. The left graph shows the baseflow and quickflow distribution based on the theory explained under section 'Delayed baseflow response in tropical Cerrados watersheds' and the graph on the right is based on the BFI filter developed by Morawietz, (1997)

Manual calibration results

Based on the hydrological knowledge the manual calibration was based more on the left graph of figure 43. Manual calibration resulted in adjustments to the paramters ESCO, CN2, SOL_K and RCHRG_DP. Figure 44 shows how the initial model simulation results change after every parameter adjustment. Showing total summed discharge, baseflow and quickflow per month in mm. These values are averaged for the entire calibration period

from 1986 until 1991. Every graph shows changes compared to the previous graph. The dotted orange line shows the previous simulated discharge and the blue line shows the observed discharge.



1. Initial values
2. ESCO is changed in order to reduce PBIAS
3. CN2 is changed in order to reduce the quickflow component and increase baseflow
4. SOL_K is reduced to reduce lateral flow and increase baseflow
5. RCHRG_DP is increased to increase the reservoir time of baseflow, reducing in a baseflow peak later in the year.

Figure 44. Figured are the results of SWAT CDI showing the average observed discharge (blue line), simulated monthly average discharge (mm/month) (orange line), simulated monthly baseflow and quickflow in respectively yellow and blue columns. The period for which the monthly values are averaged is 1986 until 1991. The figures are numbered from 1 to 5 in the upper left corner of each graph. The parameters were adjusted in this consecutive order and are described in the titel of each graph. In the upper right corner of each graph the performance indicators NSE, PBIAS and RSR are figured. The lower right cell of the table shortly describes each graph.

Between step 3 and 4 the NSE value becomes more negative. This would suggest that lowering the sol_k value does not improve discharge simulation, but this is not the case. The way in which the calibration is presented in figure 44 can be a bit misleading. An important fact to understand is that these parameters are interrelated and co-dependent. If for instance the sol_k value would be 0 no water would infiltrate into the soil. Changing the rchrg_dp parameter would have no influence on the model. If sol_k would be higher than 0, changing rchrg_dp will have an effect. The parameter adjustments calibrate the model as a 'set' of adjustments that are able to make the model produce monthly average simulated discharge values with a NSE>0.717. Table 14 shows the initial parameters and the parameters that were used for the finished regional watershed model.

Table 14. Parameter values of the regional watershed model, showing units, the initial values and the values after calibration.

Parameter	Unit	Initial value	Value in regional watershed model
SOL_AWC	mm/mm	Between 0.136 – 0.202 dependent on soil type	Between 0.136 – 0.202 dependent on soil type
CN	-	Between 61 and 78 dependent on the land-use type	0.8 * the initial values for all land-use types
GW_REVAP	-	0.02	0.02
GW_DELAY	days	31	31 days
ESCO	-	0.95	0.80
EPCO	-	1.00	1.00
RECHARGE_DP	-	0.05	0.95
ALPHA_BF	days	0.048	0.048
SOL_K	mm/hr	Between 183 and 1686 dependent on the soil type	0.8 * the initial values for all soil types

Figure 45 shows that the monthly observed discharge over the period of 1986 until 1991 can be simulated by the SWAT model with a NSE of 0.717, which can be judged as satisfactory according to D.N. Moriasi & J.G. Arnold (2007). Moriasi *et al* (2007) state that discharge simulations can be judged satisfactory if NSE>0.50 and RSR ≤ 0.70, and if PBIAS ±25%.

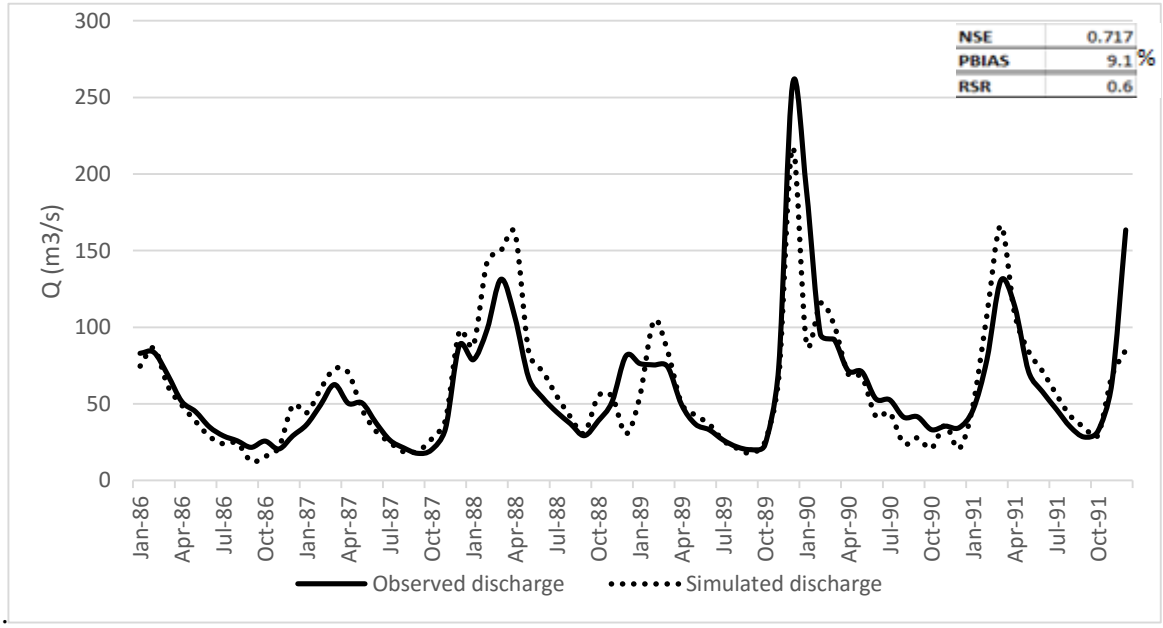


Figure 45. Monthly average observed and simulated discharge in the regional watershed after manual calibration. The black line is observed monthly average discharge (m3/s) and the dotted line is simulated discharge (m3/s). In the right upper corner the performance indicators are indicated.

11. Validation

11.1 Validation of the regional watershed model

Figure 46 shows the monthly average discharge (m3/month) results for the simulation of the validation period (1992 untill 1994). The NSE value is slightly higher than the calibration period. This possibly results from the calibration period having more variation in dry and wet years compared to the validation period. Although 1992 begins with a very wet month. 1993 and 1994 have quite average years. The higher NSE value and lower RSR value thus result from easier to simulate validation years and possibly it might have to do with more common precipitation data.

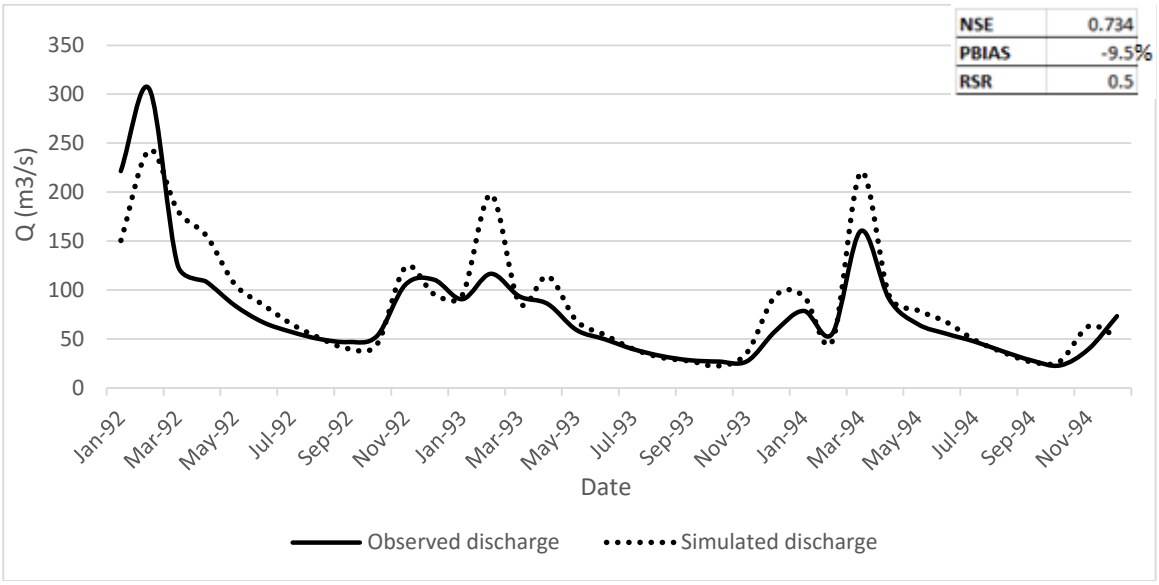


Figure 46 Average monthly observed (black line) and simulated discharge using the regional watershed model (dotted line)

11.2 Validating the downscaling methodology

In this paragraph the downscaling method described in figure 6 of chapter 5 is applied for the study area. The parameter set of the local watershed model is adjusted based on the parameter adjustments done during the calibration of the regional scale model, i.e. the parameters esco, cn2, sol_k and rchrg_dp. Figure 47 shows average monthly observed discharge from the Buriti Vermelho River during the five year validation period of 2008 until 2012. The water extracted by the canals has been calculated and added to the observed discharge at the Encontro hydrological station, which is located at the watershed outlet. The figure also shows the first discharge simulation at the watershed outlet of the local watershed model, using the initial parameter values chosen by SWAT. For esco and rchrg_dp the initial values were respectively 0.95 and 0.05. Sol_k and cn2 were dependent on respectively the soil type and the land-use type of the HRU. There were sol_k values for the most prevalent soiltypes:, cambisol and red latosol are respectively 308 mm/hr and 1986 mm/hr.

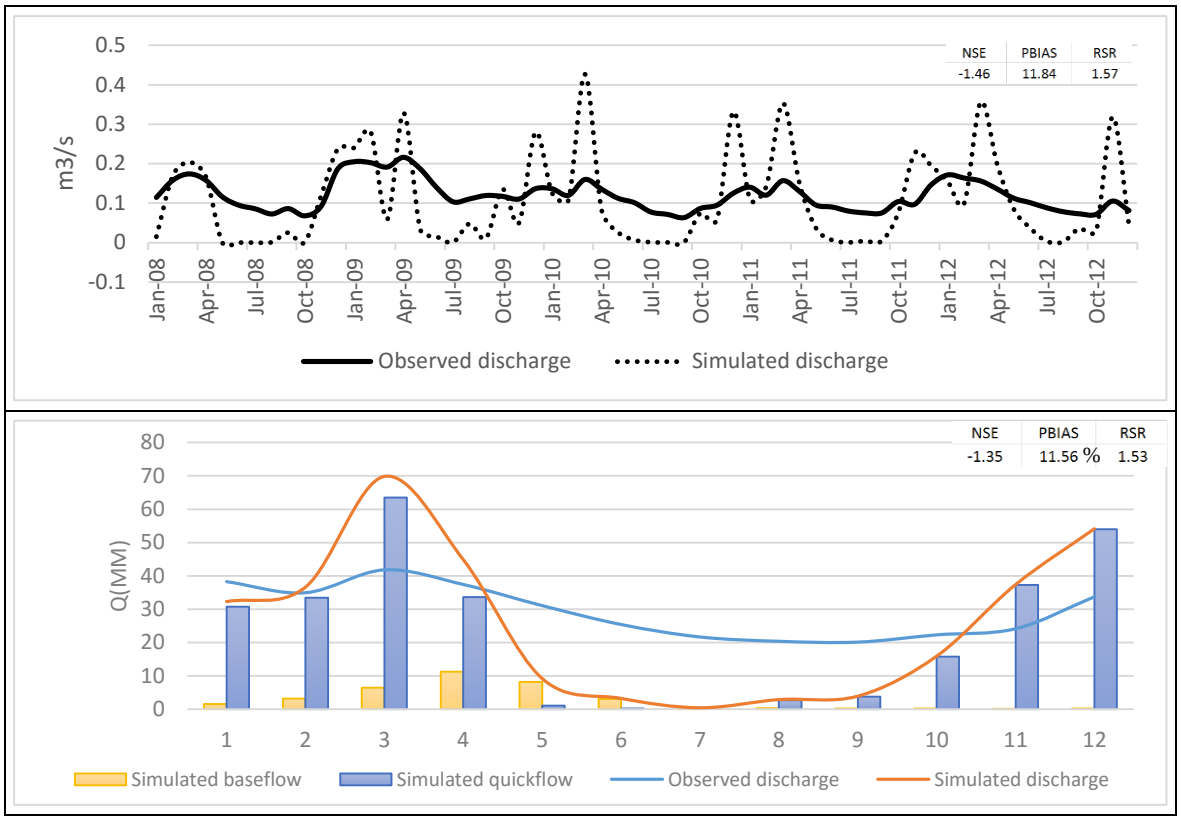


Figure 47. Local watershed model discharge simulation of the Buriti Vermelho river with initial parameter values compared to observed discharge

The NSE value for indicated in the upper corner for both time series is a negative value. The PBIAS indicates that the model is underestimating the amount of discharge. The average summed yearly discharge is simulated to be 310 mm and the observed discharge is 351 mm. A much higher percentage of the discharge is determined by quickflow than baseflow and in the dry periods the model does not predict any flow. Approximately the same issues as during calibration of the regional scale model. Figure 48 shows the simulated discharge using the parameter adjustments from the regional scale model compared to observed values.

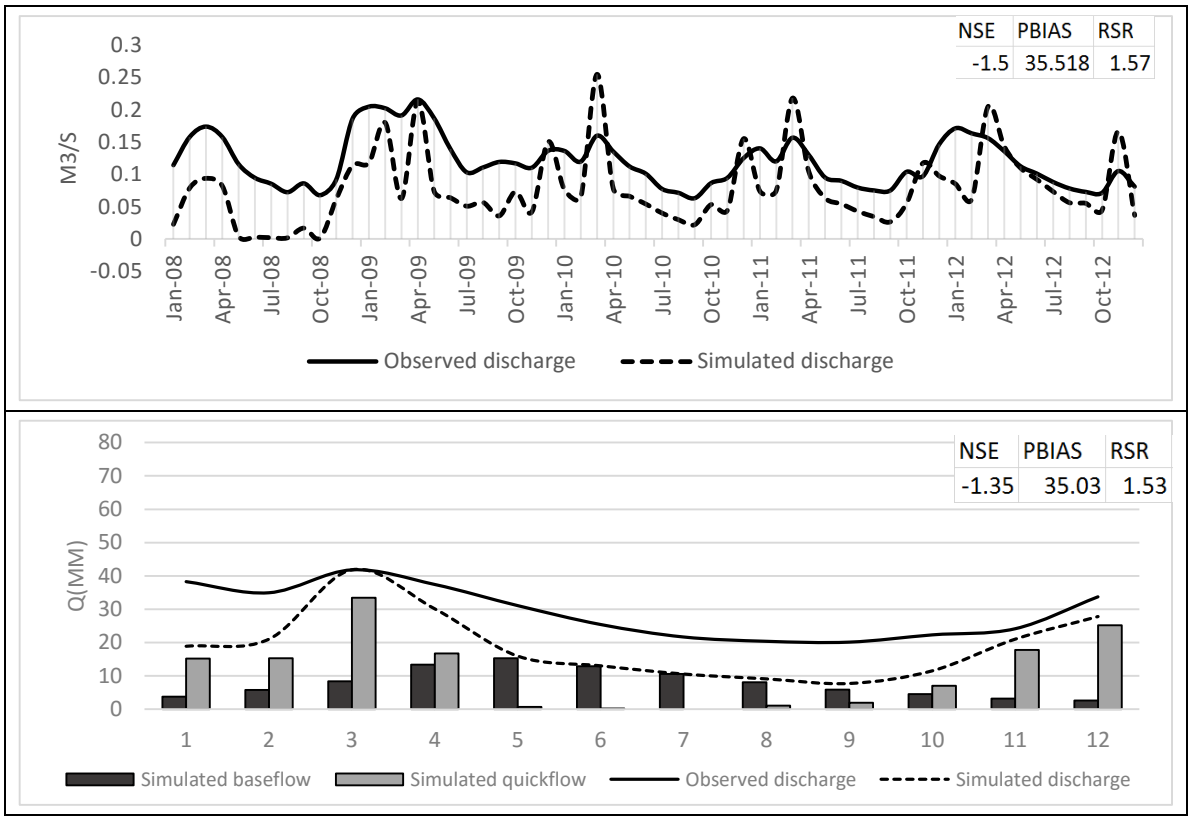


Figure 48. Local watershed model discharge simulation of the Buriti Vermelho river using adjusted parameter values from the regional scale model compared to observed values

From figure 48 we can see that the baseflow and quickflow distribution changed. Increasing the flow component baseflow. Moreover the discharge in most years is not zero during the dry period, except for 2008. Sadly the NSE value stayed around the same value and did not really increase the prediction performance. The most important change compared to the initial simulation is the increase in PBIAS, which means that the model is severely underestimating the observed discharge.

The main parameter adjustment resulting in a large change of total water yield in the river results from adjusting esco, which can be concluded after reading the chapter about sensitivity analysis. Of the four parameters only esco determines a change in discharge amount. This was the main function of the parameter during manual calibration. After each parameter adjustment, the PBIAS value was checked in order to see if the total discharge was not over or underestimated. If this was the case the esco value was adjusted. The final value for the regional model was 0.75. The other parameters determined the distribution of water amongst the different reservoirs in the SWAT model; flow over the surface, flow in the soil layer, flow in the shallow aquifer and flow in the deeper aquifers, and thus also determined the reservoir time of the water. Reducing sol_k increased the reservoir time of water in the soil layer and increased percolation to the shallow aquifer. Reducing CN2 meant more water infiltrated in the soil, which reduced surface runoff and increased lateral flow and baseflow. Increasing rchrg_dp increased the percolation to the deeper aquifer, resulting in a longer groundwater delay time (time it takes for water to reach the river). In conclusion the three paramters sol_k, cn2 and rchrg_dp determine the distribution between quickflow and baseflow and determine the reservoir time of baseflow. Another way to put it is that sol_k, cn2 and rchrg_dp determine the shape of the baseflow, quickflow columns and discharge line chart in figure 48, while esco determines the magnitude of the of discharge. Figure 49 illustrates this.

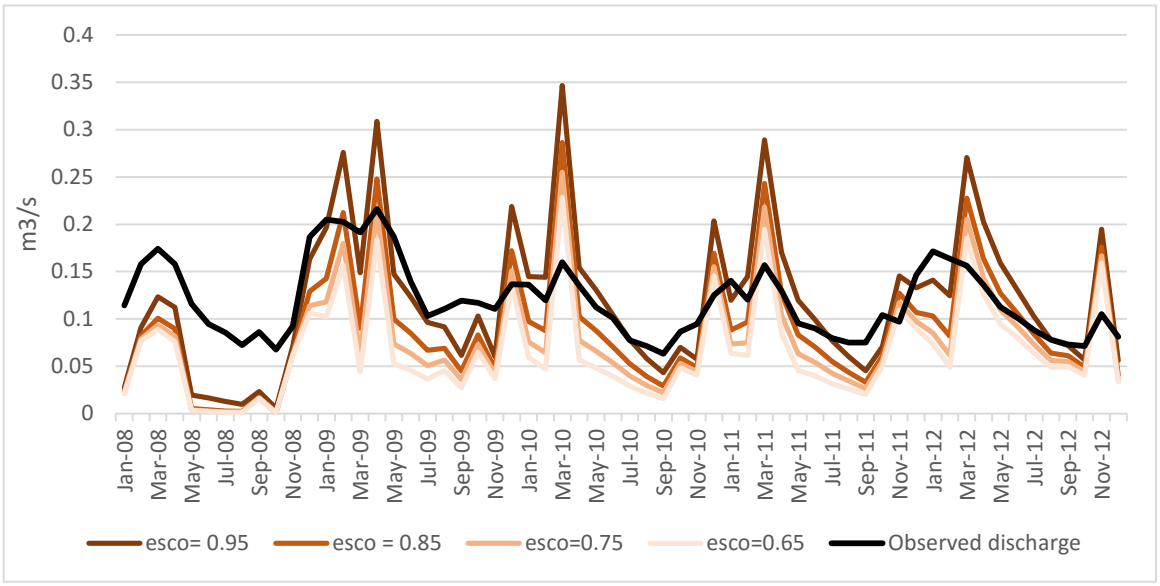


Figure 49. Esco adjustments on the monthly average discharge simulation

As sol_k, cn2 and rchrg_dp are the calibration parameters which determine the shape of the monthly average discharge amounts. These parameter adjustments play an important role in correcting the baseflow/quickflow distribution and the baseflow reservoir time. Adjusting esco is only important to attain the correct discharge amount. The esco parameter adjustment cannot be based on the parameter adjustment done in the regional watershed model and must be based on other information. Equation 22 indicates that in a catchment under stationary state conditions precipitation equals the sum of discharge and evapotranspiration. The storage term dS/dt becomes zero.

$$P = Q + ET \qquad (22)$$

The regional watershed contains data about these three terms. Data is acquired about precipitation and discharge and evapotranspiration is determined using Penman-Monteith. As the local watershed is inside the regional watershed the ratio between evapotranspiration and discharge should be approximately the same, because averaged over a longer time period the weather conditions and precipitation amounts are approximately the same per square meter. In order to estimate the total water yield the information about the ratio between discharge and precipitation can be used to estimate the amount of discharge in the local watershed. This assumption is summarized by the equation 23:

$$\frac{\sum Q_{regional\ watershed}}{\sum P_{regional\ watershed}} = \frac{\sum Q_{local\ watershed}}{\sum P_{local\ watershed}} \qquad (23)$$

The $\sum$ operator takes the sum of the daily variable values during the time series period. For the regional model both $\sum Q$ and $\sum P$ are known. In an ungauged local watershed $\sum P$ is known, but $\sum Q$ is unknown. Using equation 23. $\sum Q$ can be estimated. After estimation of $\sum Q$ for the local watershed the esco parameter can be adjusted accordingly to attain this value.

To test the validity of this theory the sum of precipitation and discharge of the Preto river and the Buriti Vermelho river at the watershed outlet is calculated. The precipitation time series are in mm, but the discharge time series in m^3/day. The observed discharge values need to be multiplied by the watershed area and the days in the summation period and finally multiplied by 1000 to change meters in to millimeters. This results in the total discharge in mm. The calculated sums and the calculated Q-P ratio for the regional and the local watershed are shown in table 15.

Table 15. Values calculated based on equation 23. Summed values for the daily discharge and precipitation in mm for the Preto and Buriti Vermelho river. The column on the right shows the ratio between the summed discharge and precipitation.

	Summation period	(mm)	$\frac{\sum Q}{\sum P}$
$\sum Q_{Rio\ Preto}$	1986 -1994	4428	0.3649
$\sum P_{regional\ watershed}$	1986 -1994	12132	
$\sum Q_{Buriti\ Vermelho}$	2008 -2012	1755	0.3702
$\sum P_{Local\ watershed}$	2008 -2012	4740	

The resulting ratios differ 0.0053 from each other, thus confirming the assumption made in equation 17 is justifiable. In the case of Buriti Vermelho the discharge was measured. The goal of this downscaling methodology is to assess rivers of which the discharge is not measured. Using equation 17 to estimate $\sum Q_{local\ watershed}$ and then adjusting the parameter esco so the SWAT model simulates in the correct order of magnitude.

Based on equation 23 the downscaling methodology has been adjusted. It now consists of two steps.

Adjust parameters that determine the dynamics of a watershed based on the parameter adjustments made in a regional watershed, namely sol_k, cn2 and rchrg_dp.

Adjust the parameter esco that determines the ratio between Q and P, based on time series analysis of using equation 17.

The final discharge simulation and the calculated performance indicators are presented in the results Chapter III.

Chapter III

Results and synthesis

12. Results

12.1 Results and synthesis of downscaling methodology

After adjustment of the parameters of the local watershed SWAT model according to the two steps described in 11.2 two graphs can be generated. Figure 50 shows monthly average observed discharge at the Buriti Vermelho river and simulated discharge (orange) using the local watershed model, which was calibrated using the downscaling methodology. During the wet season the model overestimates the discharge amount. The NSE valus is negative indicating an unsatisfactory model simulation. The PBIAS value is however less than 1%, which is a good result.

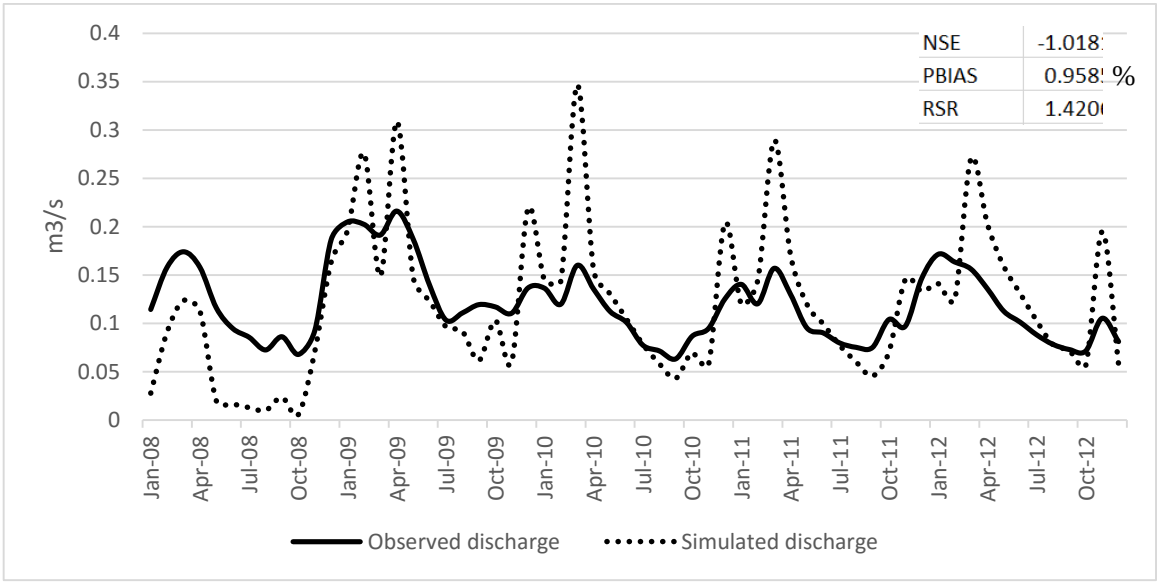


Figure 50. Monthly average observed discharge at the Buriti Vermelho river (black line) and simulated discharge (Dotted line) using the local watershed model, which was calibrated using the adjusted downscaling methodology described in Chapter 11.2

The monthly average simulated discharge over the entire validation period is shown in figure 51 as well as the simulated baseflow and quickflow. Performance indicators are indicated in the right upper corner.

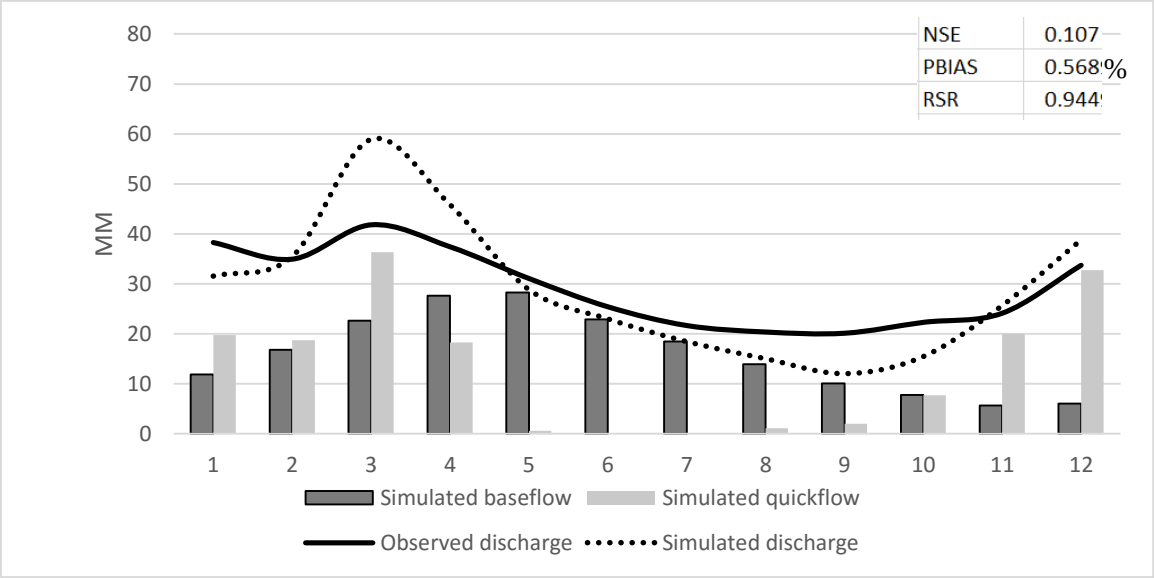


Figure 51. Monthly average observed discharge (black line) and simulated discharge (dotted line) using the local watershed model, which was calibrated using the downscaling methodology in described in 11.2. The dark grey column is simulated baseflow and the light grey column is simulated quickflow.

Synthesis of downscaling methodology results

Figure 50 and 51 show the results of the local watershed model calibrated using the regional watershed parameter adjustments. Figure 50 shows the monthly discharge simulation compared to the observed discharge over the period January 2008 until December 2012, of which can be concluded that the model underestimates peaks during wet seasons. The recession curves are approximately at the right times as the observed recession curves and seem to fit moderately well. During the dry months the model underestimates the discharge. This monthly discharge prediction has a negative NSE value of -1.08, PBIAS of 0.96% and a RSR of 1.42. According to the criteria by Moriasi *et al* (2007) to assess goodness-of-fit, the local watershed model simulations are unsatisfactory according to the NSE and RSR performance indicators. The NSE value is less than zero and thus according to the NSE performance indicator the prediction is worse than the average of the observed discharge time series. The PBIAS value on the other hand is very low, with a value of less than 1 percent. The closer to zero the better the model fit. This means that the model estimates the discharge of the Buriti Vermelho in the right order of magnitude. If the goal of model use were to estimate roughly what the discharge is in a small watershed is, then the methodology could result in a usefull prediction. The monthly dynamic of the simulated discharge should be considered with a very high level of uncertainty, especially during the wet months.

Figure 51 shows the output graph averaged over the entire period. This twelve month time series has a NSE of 0.107, which is a positive NSE value and thus a better predictor than the average of the observed discharge time series. As most of these local watersheds are ungauged, a prediction of discharge which is only just better than the average of the observed value can already be considered as valuable. The local watershed model, could for instance be used to predict how climate change or land-use change would impact the average discharge of the Buriti Vermelho River. If the model predicts only a few percentages change due to some environmental impact, this should not be considered as a result with a high certainty. Alternatively, if the model predicts that the average discharge changes drastically with a percentage change of more than 25%, this result exceeds the uncertainty bands of the model prediction and should be taken carefully.

12.2 Results and synthesis of land-use change impact assessment

The land-use scenario analysis done with the regional watershed model is presented in figures 52, 53 and 54. Figure 52 and 53 show the results of the simulations of monthly average discharge, effective evapotranspiration, baseflow and quickflow. The black line indicates a simulation done with the regional watershed model based on the land-use map of 1984. The dotted line indicates a simulation with the 2014 land-use map.

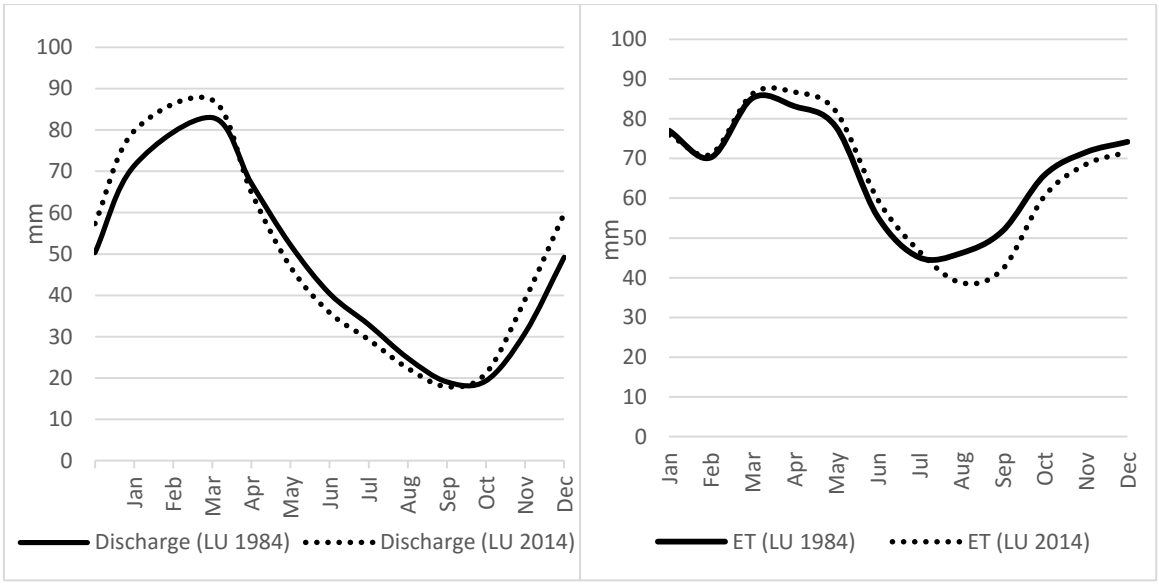


Figure 52. Monthly average values in mm for discharge (left) and effective evapotranspiration (right). The black line indicates a simulation done with the regional watershed model based on the 1984 land-use map and the dotted line with the 2014 land-use map.

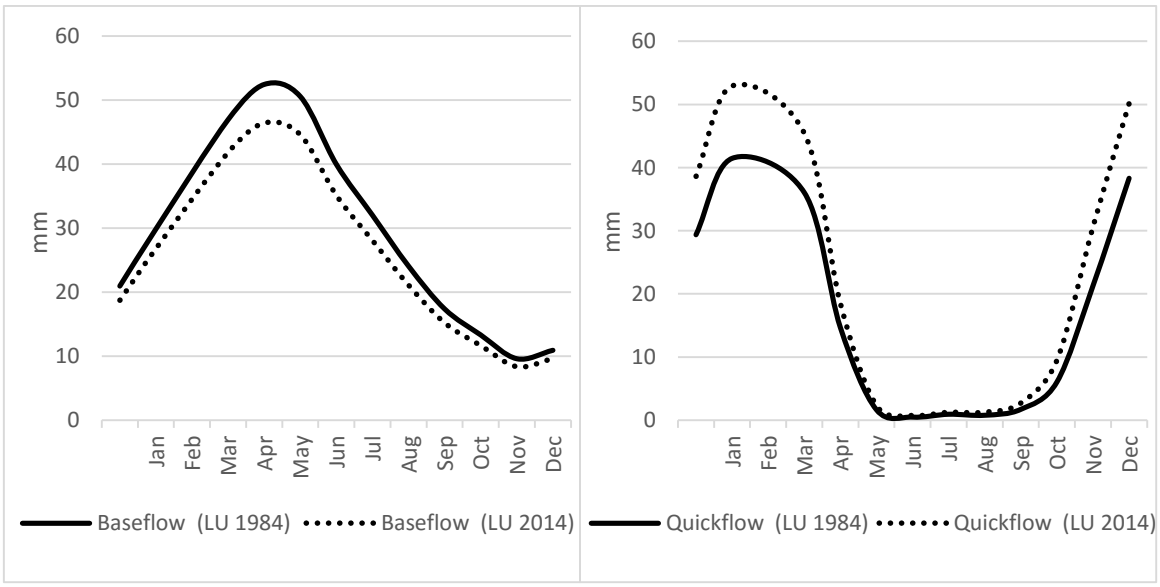


Figure 53 Monthly average values in mm for baseflow (left) and quickflow (right). The black line indicates a simulation done with the regional watershed model based on the 1984 land-use map and the dotted line with the 2014 land-use map.

The weather data during this period is kept constant. It shows how land-use impacts components of the water balance and flow components.From the upper left total discharge graph we can see that 2014 discharge is lower in the wet period (Jan – april) and higher in the dry period compared to the discharge with the 1984 land-use map. More water thus infiltrates in the dry season. From the upper right graph we can see that evapotranspiration is higher during the wet period months of April and May, but much lower in the drier period from August until December. With the land-use map of 2014 thus less evapotranspiration occurs

during periods of low water availability. From the lower two graphs we can see that baseflow decreases and quickflow increases using the 2014 land-use map compared to the 1986 land-use map.

Figure 54 shows what these changes look like for a selection of the simulated time series, namely from January 1991 until December 1994. A general trend showing a decrease in discharge is noticeable resulting from the change of land-use maps. During this period the increased discharge during the months January, February and March shown in the left graph in figure 52 is not recognizable in figure 54. Only in the year 1991 is the discharge simulated using the 2014 land-use map higher than the discharge simulated with the 1984 land-use map.

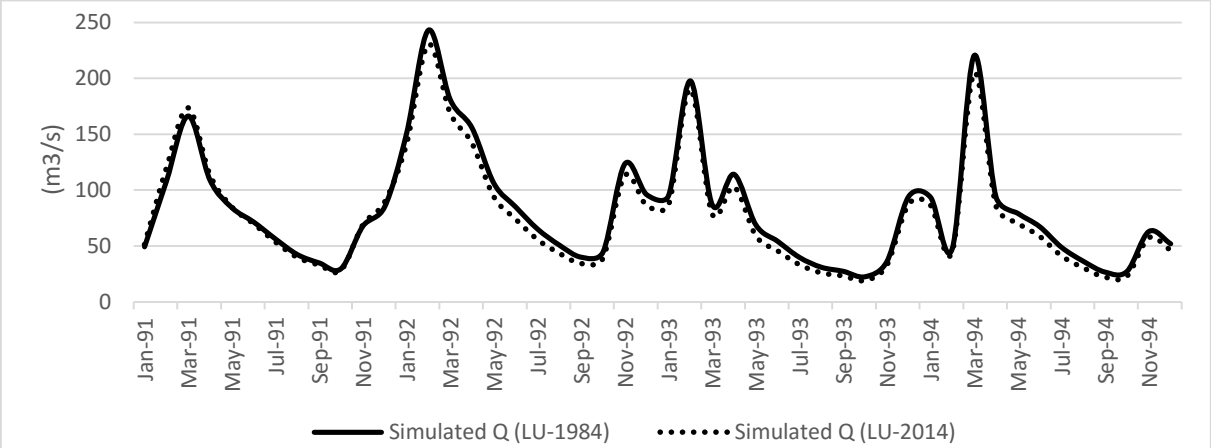


Figure 54. Simulated monthly average discharges using land-use maps from 1984 and 2014 as input for the Regional Watershed model

Synthesis land-use change impact results

Land-use change impact was assessed by simulating discharge using the calibrated and validated regional watershed model, which gave an NSE value of 0.734, PBIAS of -9.5% RSR of 0.5. The criteria of Moriasi *et al.* (2007) indicate that the model is satisfactory when: NSE>0.5, RSR<1 and PBIAS< ±25%. The regional model thus predicts with a high level of certainty. The model first simulated discharge during the period 1986 and 1994 using the land-use map of 1984. Then the land-use map was exchanged for the land-use map of 2014. The methodology is explained in detail in 6.2.

This method isolates land-use as variable impacting discharge. Observed discharge in the period 1984 until 1994 could also simply be compared to observed discharge between 2004 and 2015 and conclusions could then be drawn about the impact of land-use change. The result from this analysis however does not single out 'land-use change' as the only environmental change. One important environmental change is for instance the Queimedos dam, which in the case of this later methodology is not taken into account.

The results of both simulations can be viewed in figure 54. From the graph can be observed that discharge simulation using the 2014 land-use was lower compared to discharge simulated using the 1986 land-use map. The area of savannah severely decreased during this period with 18.4% and agricultural land increased with 23.1%. The lower discharge value could be explained by the fact that in SWAT the land-use type agriculture has a higher evaporative demand compared to the land-use type savannah.

The SWAT model also gives information about baseflow, quickflow and evapotranspiration. For both land-use simulations these outputs are shown in figure 52 and 53, which show the monthly sum of discharge, evapotranspiration (ET), baseflow and quickflow averaged over the whole simulation period (1986 until 1994). From the discharge graph we can conclude that in general discharge simulated with the 2014 land-use map is higher in the months November until April, i.e the wet season and discharge is lower during April and September, i.e. the dry season. Evapotranspiration is higher using the 2014 land-use map during the wet season and lower during the dry season, compared to evapotranspiration using the 1984 land-use map. Baseflow simulated using the 2014 land-use map is lower during the whole period and quickflow is higher in the wet season, both compared to simulated discharge using the 1986 land-use map. Especially in the months December until March quickflow is much higher. The main reason for this is that the curve number, which determines the ratio between infiltration and overland flow, is higher for the land-use type agriculture

compared to savannah. A higher curve number means less infiltration and more overland flow, i.e. more quickflow and less baseflow. So the model simulates this accurately.

The main observations that can be made is that the drastic land-use change of savannah decreasing and agriculture increasing has led to a decrease in discharge. Evapotranspiration is lower the dry season. Baseflow has become lower and quickflow higher, due to agriculture having a higher curve number than savannah. Continued decrease in savannah and increase in agriculture in this region will lead to lower baseflow in the Preto River and a total decrease in discharge.

12.3 Results and synthesis of climate change impact assessment

Figure 55 shows the results of the discharge simulation using the regional watershed model runs using rcp 4.5 scenario's for the period 2050 (orange line) and 2100 (red line). The figure also shows the simulation with no changes to the climate data (green line). The weather data that was adjusted was from the period 1986 until 1994.

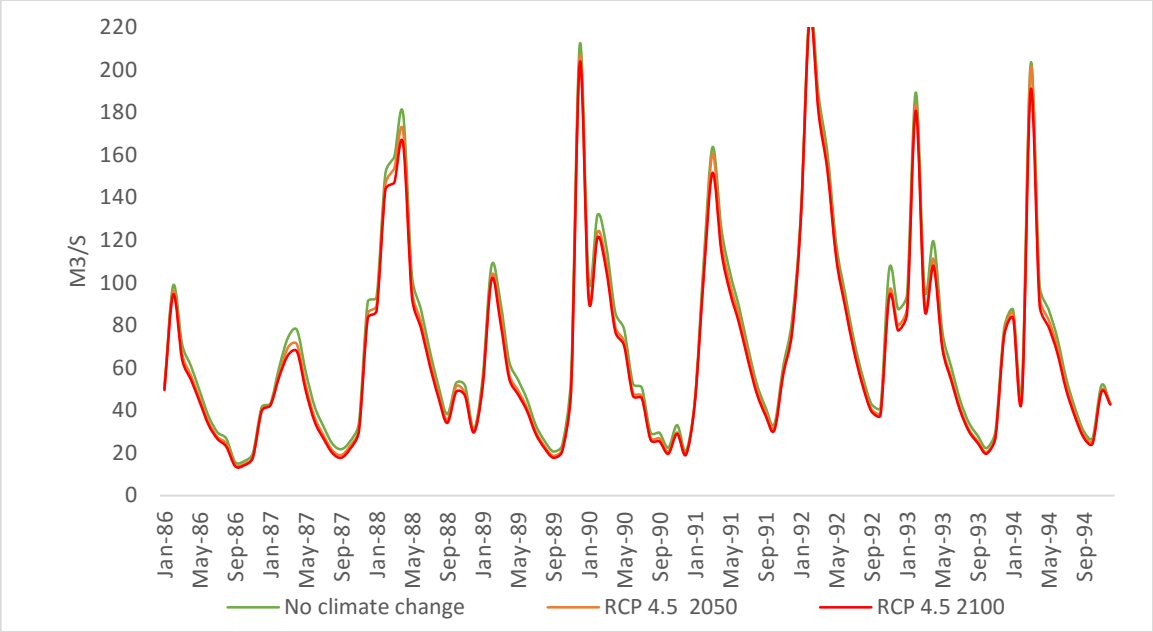


Figure 55. RCP4.5 scenario discharge predictions using the regional watershed model

In figure 56 the results are averaged over the entire period to produce monthly average sum values. Baseflow and quickflow output is also summarized in this graph. The green line is the observed average monthly sum of discharge without any changes to the climate input data.

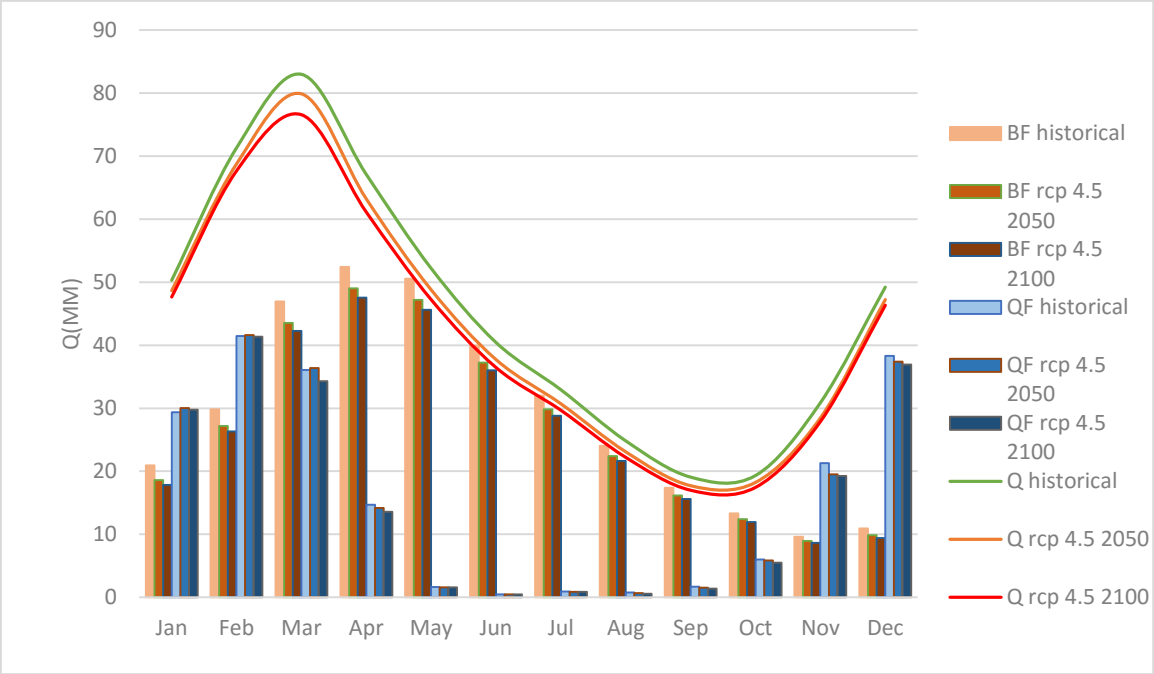


Figure 56. Observed and simulated monthly average discharge over the period 1986 until 1994 until 1994, predicted using the rcp 4.5 scenario in the regional watershed model

Figure 57 and 58 shows the same results as presented in figure 55 and 56, but now using the RCP 8.5 scenario weather data as input.

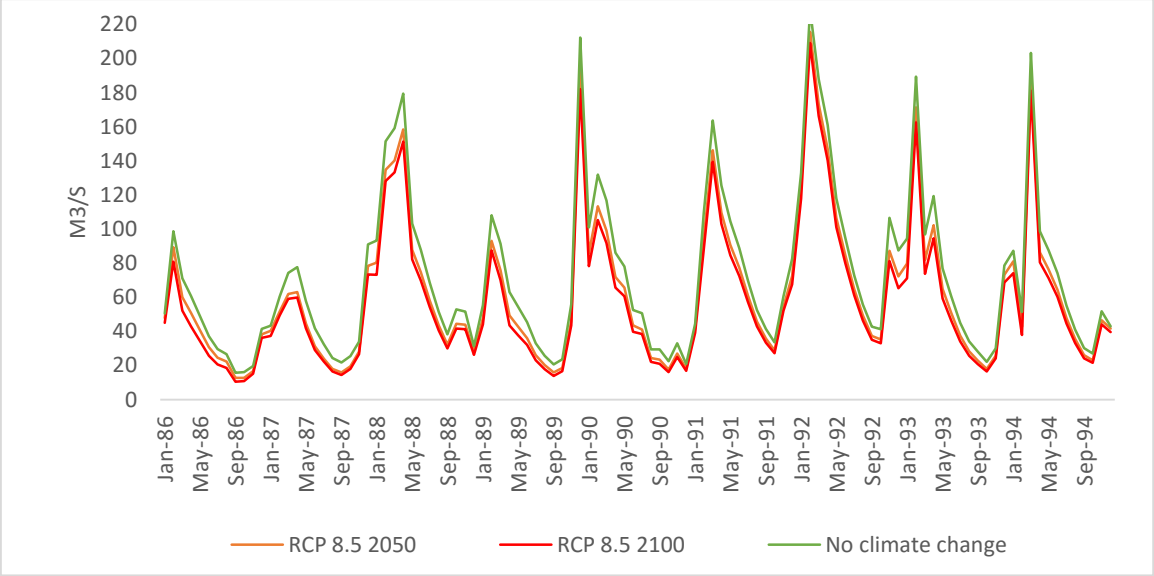


Figure 57. RCP4.5 scenario discharge predictions using the regional watershed model

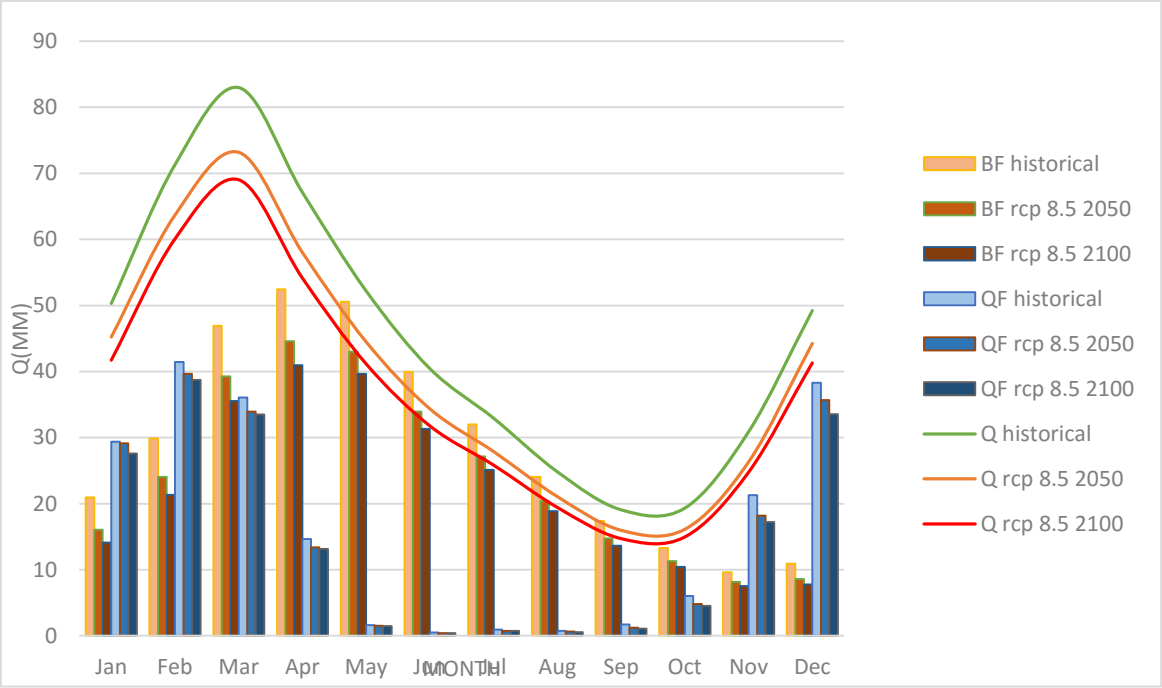


Figure 58. Observed and simulated monthly average discharge over the period 1986 until 1994 until 1994, predicted using the RCP8.5 scenario in the regional watershed model

Table 16 shows percentage change in baseflow, quickflow and discharge resulting from the rcp scenario weather input simulations using the regional watershed model.

Table 16. Percentage change compared to historical average

IPCC climate change scenarios	Year	Baseflow	Quickflow	Discharge
rcp 4.5 2050	2050	-7.3%	-1.4%	-5.2%
rcp 4.5 2100	2100	-10.4%	-3.7%	-8.0%
rcp 8.5 2050	2050	-16.2%	-6.9%	-12.9%
rcp 8.5 2100	2100	-23.4%	-5%	-18.8%

Synthesis climate change impact results

Two RCP climate change scenarios, RCP4.5 and RCP8.5, have been used as input for the regional watershed model to predict the impact of temperature and precipitation change on the Preto River. Figure 55 and 57 show the discharge simulation over the entire output period, from 1986 until 1994. The simulation with no precipitation or temperature adjustment has a goodness-of-fit of NSE = 0.734 compared to observed discharge data. From the graphs can be concluded that discharge decreases according to these rcp scenarios, with rcp8.5 giving more drastic decreases compared to RCP4.5. Moreover the adjustments based on 2100 lead to a higher decrease in discharge for both RCP scenarios. Figure 56 and 58 show the monthly sum of discharge averaged over the simulation period. Results are shown for the RCP4.5 and RCP8.5 scenario for monthly discharge values, baseflow and quickflow. In the RCP4.5 scenario baseflow decreases from 2050 until 2100 compared to the historical discharge simulation. Quickflow decreases too, except in the month January for the RCP4.5 scenario. For the RCP8.5 scenario both quickflow and baseflow decrease from 2050 untill 2100. The decrease is quite drastic for this worst-case IPCC scenario.

Table 16 shows the percentage change in baseflow, quickflow and discharge compared to the historical discharge simulation. In the worst case rcp 8.5 scenario the change in 2100 in discharge is -18.8 percent. As the average discharge in the Preto river is approximately 50 m^3/s, this means a decrease of 10m^3/s, which means that in a day 8.64 * 10^5 m^3 of water flows less through the river. The main flow component that changes is baseflow, which in 2050 for RCP4.5 is predicted to change with -7.37% and for RCP8.5 with -

10.4% compared to the simulated historical baseflow. In all the scenarios quickflow decreases as well, but not as much as baseflow.

In conclusion the scenario's all result in a decrease in discharge in the Preto river. In the case of RCP4.5 in 2050 the average discharge decreases with -5.2% and for RCP8.5 this is -12.95%. In 2100 for RCP4.5 the change is expected to be -8% and for RCP8.5 it is -18.8% compared to historical discharge data. Lastly in both scenarios baseflow decreases more drastically than quickflow.

12.4 Results and synthesis of Queimados dam impact assessment

Figure 59 showing observed discharge from January 2009 until November 2012 and simulated monthly average discharge using the regional watershed model. Figure 60 shows monthly average values of the observed and simulated discharge averaged over the four year period.

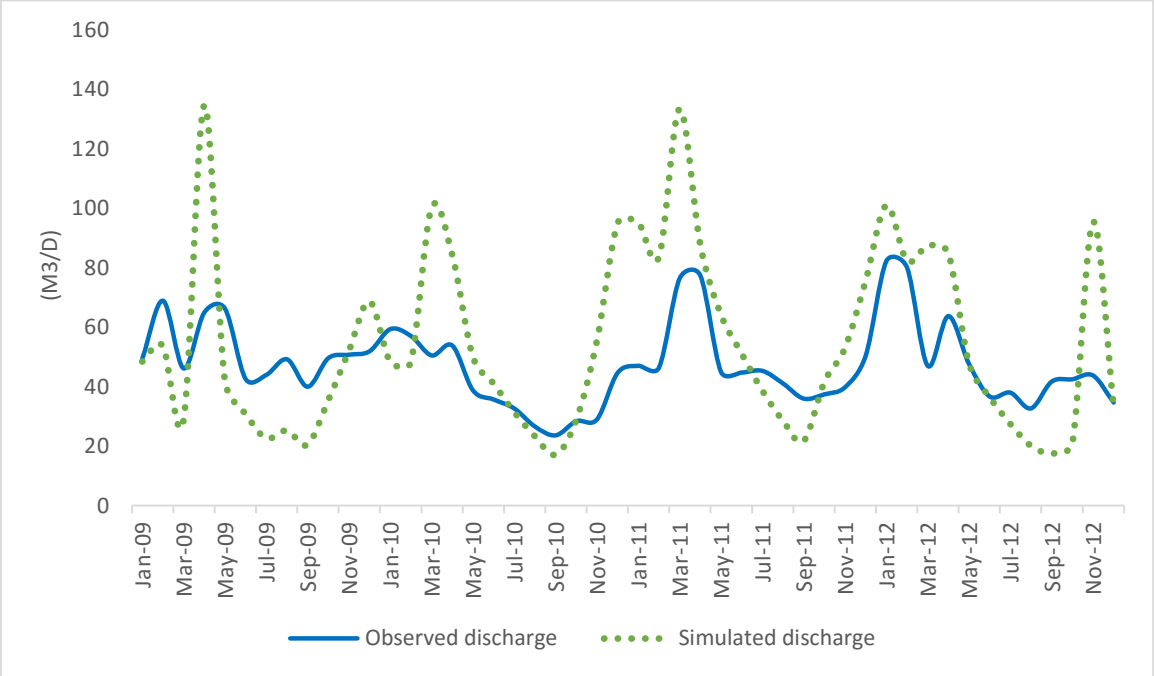


Figure 59. Observed and simulated discharge using the regional watershed model. The simulated discharge represents natural discharge, without the influence of the hydro-electric dam.

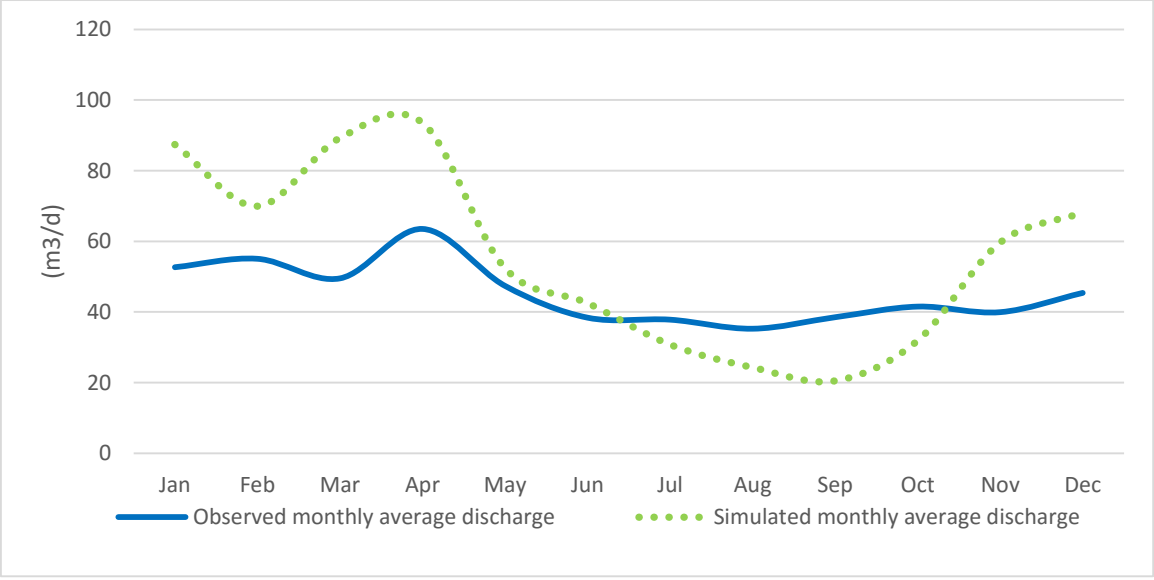


Figure 60. Observed and simulated discharge using the regional watershed model averaged over the entire simulation period. The simulated discharge represents natural discharge, without the influence of the hydro-electric dam.

Synthesis Hydropower dam impact results

To determine the impact of the Queimados dam the regional watershed model simulated discharge based on the land-use map from 2014 and weather data between 2004 and 2014. Figure 59 shows the simulated discharge during this period compared to the observed discharge. The results show that Queimados dam has two large impacts. Firstly the amount of discharge is 16.4% less over the period 2004 until 2014. This is probably due to the irrigation canals that extract water from the Queimados reservoir. A second observation is that seasonal variation in the simulated discharge is much higher than the observed discharge. This is also not an unexpected result. Management of the hydro-electric dam and a controlled steady water level of the Queimados reservoir result in observed discharge in the Preto river showing less seasonal variation. This result is clearly shown in figure 60, which shows the average monthly simulated and observed discharge.

Chapter IV

Conclusions

This thesis has attempted to provide new (hydrological) insights in the challenges that come with the rapid population growth, increased use of cropland and the construction of a dam in the Federal District in Brazil. This thesis presents two hydrological SWAT models, one of a regional watershed belonging to the Preto river and a second one of a smaller local watershed of the Buriti Vermelho river.

Firstly, a downscaling methodology was tested. As was argued, a downscaling tool can be of great use to predict influences of (climate) change on discharge, when discharge data of a larger regional watershed is known, but data on the more vulnerable smaller rivers is lacking. The discharge data from the Preto river was used to simulate the discharge of a smaller river with the SWAT model. On a low temporal resolution the downscaling methodology results in discharge simulations that correspond to the average observed values with an NSE > 0.1. On a monthly temporal resolution the methodology predicts discharge worse than the average of the observed discharge. This means that the methodology provides unsatisfactory results for analysis of ungauged small rivers in the Cerrados.

Secondly, the regional watershed model of the Preto river was also used to assess the impacts of various environmental changes, namely climate change, land-use change and a hydropower dam on discharge. The main impacts on the discharge of the Preto river are a significant decrease of baseflow and a slight increase in quickflow. Both RCP4.5 and RCP8.5 scenarios result in a trend of decreasing total discharge from 2050 until 2100. In the RCP8.5 scenario run for 2100, there is a 18% decrease in total discharge. Moreover, baseflow responds more to the climate changes than quickflow. The hydro-electric dam severely impacts the Preto river in two ways. Firstly, average yearly discharge decreases with 16.4% and seasonal variation in discharge is decreased. In addition, peaks in wet seasons are much lower and during dry periods more discharge is simulated. The scenario analysis shows that the three environmental impacts severly impact the hydrology of the watershed of the Rio Preto river, but in different ways. Climate change and land-use change mainly impact baseflow and the hydropower dam mainly influence seasonal fluctuations in baseflow. All of them however decrease the total yearly amount of discharge.

14. Discussion

Calibration

Calibration of the regional watershed model was done manually which resulted in satisfactory simulation results. However in chapter 9.3 the SWAT-CUP program was described, which uses various calibration methods to automatically calibrate watershed models. Several attempts were made to calibrate the regional watershed model automatically using this program. Parameter ranges were determined around the most sensitive parameters. The scripts inside SWAT-CUP were written and seemed to be correct, but unfortunately automatic calibration did not work. Using this automatic method to further enhance the calibration would be an improvement to the certainty of the model predictions.

Downscaling methodology

The downscaling methodology produced output that could be used on a very low temporal resolution, i.e. average monthly discharge in a period of multiple years. The generated output gives discharge simulation values which are better than the average value of the observed discharge in small rivers. The goodness-of-fit of the local watershed model increased significantly after applying the parameter adjustments based on the regional watershed model. One of the main challenges was to attain a discharge simulation during the dry season. The downscaling methodology adjusted the output to the extent that discharge in the dry season was simulated, i.e. the local watershed model was able to capture the seasonal variation of discharge.

Reasons that prevented good results are probably due to have several reasons. Firstly, the local watershed location is influenced more by spatial variation in precipitation, because of the smaller area and the specific elevation compared to the regional watershed. As the regional watershed is much larger, the spatial variation is smaller. Moreover, as precipitation is correlated with elevation, a difference in average elevation between

the local and regional watershed can also be problematic when applying the downscaling methodology. Secondly, the ratio of savanna to agriculture as well as the distribution of soil classes is very important in how the water balance in the SWAT models is determined and thus for the calibration of the model. As the local watershed had a higher percentage of agricultural land compared to the regional watershed and the soil distribution is dissimilar, calibration for both model cannot be done in the same way.

Input data

Calibration of a model is highly dependent on the input data, which is a very important notion. A large portion of the input data was similar for the local and the regional watershed, such as weather parameters, soil types, precipitation amounts and seasonality. Nevertheless, as these parameter values had still differences, this resulted in prediction performance that was useful on a very large temporal timescale.

A methodology adjustment would be to select the regional watershed based on the most SWAT input similarities with the local watershed model. This can be done by determining a watershed boundary of the regional watershed based on percentage distribution of soil types, land-use types and average elevation of the local watershed. If the ratio of soil types and the ratio of land-use types and the average elevation of the regional watershed is similar to that of the local watershed, this would probably significantly improve the results of the downscaling methodology. This was already done partly by choosing the upstream area of the Preto river watershed instead of using the full watershed for the regional watershed model. The reason for this choice was that there was a large difference in the elevation of the upstream watershed area compared to the downstream area.

Scenario analysis

The land-use change assessment was performed with a land-use map that had a relatively low information density. The regional watershed has much more variation in land-use types. For instance the area that was modelled as agriculture could be divided into several different land-use types, incorporating irrigation practices and different crop types. This would make the regional watershed more realistic. At the same time this makes construction of the model more time consuming.

Climate change scenarios were based on projected changes in precipitation and average temperature which were based on the mean of 20 global climate models. The regional watershed model could also be run with a selection of models showing the lowest bias. A second adjustment to the methodology could be to assess more RCP scenario's such as the RCP2.6 and RCP6.5 scenarios. Finally, the SWAT model also provides the possibility to adjust carbon dioxide concentrations in the atmosphere as well as relative humidity. Since there was no data for these parameters, these could not be included in the models presented in this thesis. Adjusting these parameters according to climate model simulations should improve the climate change scenario analysis.

Improvements to the methodology of the hydropower impact assessment, apart from using a more detailed land-use map, are hard to determine. It is however interesting to investigate which processes related to the Queimados dam and the Queimados reservoir impact the discharge the most. Satellite imagery shows that a multitude of canals are connected to the Queimados reservoir. Probably water is extracted for several other reasons as well.

Possible continuation research

The regional and local watershed models could be significantly improved if more information was collected about the groundwater system. This could significantly improve calibration results as parameter adjustments could be reinforced by collected information or observed data. Examples of collected information could be time series analysis of the groundwater level from several well locations in the watershed. Moreover, as information is acquired about the geology assumptions can better be made about the movement of water in the groundwater system. Based on the found data and information a groundwater model could be constructed in e.g. MODFLOW, which could then be coupled to the SWAT model. The groundwater model would functioning as the groundwater reservoir of the SWAT model, which would simulate groundwater flow more realistically.

An attempt was already made to calibrate the local watershed model by using the Buriti Vermelho river discharge data. This deemed to be more difficult than the regional watershed model, but as a continuation research it would be very interesting to calibrate the local watershed model further until the model simulates satisfactory output, i.e. simulated discharge time series that have a goodness-of-fit with the observed discharge time series. After successful calibration the local watershed model could be used to assess

environmental impacts, such as the ones assessed in this thesis, but then on a small river. A comparison could be made with the results of this research to make conclusions about the way in which large rivers in the Cerrados respond differently to environmental changes compared to small rivers.

References

Abbaspour, K. C. (2015). SWAT-CUP: SWAT Calibration and Uncertainty Programs - A User Manual.

Abbaspour, K. C. (2016). *Calibration of Hydrologic Models: When is a Model Calibrated?*

Ahmad, H., Jamieson, R., Havard, P., Madani, A., & Zaman, Q. (2009). Evaluation of SWAT for a small watershed in Eastern Canada. *EFITA Conference*, 337–343.

Alder, J., Hostetler, S., & Williams, D. (2013). An Interactive Web Application for Visualizing Climate Data. *Eos, Transactions American Geophysical Union*, *94*(22), 197–198. http://doi.org/10.1002/2013EO220001

Arnold, J. G., Allen, P. M., Muttiah, R., & Bernhardt, G. (1995). Baseflow Program. *Ground Water*, *33*(6), 1010–1018.

Arnold, J. G., Kiniry, J. R., Srinivasan, R., Williams, J. R., Haney, E. B., & Neitsch, S. L. (2012). Soil & Water Assessment Tool: Input/output documentation. version 2012. *Texas Water Resources Institute, TR-439*, 650. Retrieved from http://swat.tamu.edu/media/69296/SWAT-IO-Documentation-2012.pdf

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., … Jha, M. K. (2012). Swat: Model Use, Calibration, and Validation. *Asabe*, *55*(4), 1491–1508.

Bressiani, D. D. A., Gassman, P. W., Fernandes, J. G., Hamilton, L., Garbossa, P., Srinivasan, R., … Mendiondo, E. M. (2015). A review of Soil and Water Assessment Tool (SWAT) applications in Brazil : challenges and prospects, *8*(3). http://doi.org/10.3965/j.ijabe.20150803.1765

Brouwer, C., Heibloem, M., & Division, D. (1986). Irrigation Water Management : Irrigation Water Needs. *Training*, *7*, 225–240. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20980772

Campozano, L., Sánchez, E., Aviles, a, & Samaniego, E. (2014). Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes. *Maskana*, *5*(1), 99–115. Retrieved from http://dspace.ucuenca.edu.ec:8080/handle/123456789/5586

de Souza Martins, E. (2009). *Compartimentacao Geomorfologica da Bacia Hidrografica do Rio Buriti Vermelho*. Planaltine.

Delvaux, B., & Brahy, V. (2016). Mineral Soils conditioned by a Wet (Sub) Tropical Climate. Retrieved April 30, 2016, from http://www.fao.org/docrep/003/y1899e/y1899e08a.htm

Embrapa. (2016). Hydrological, soil and land-use data. Sobradinho.

Feyereisen, G. G. W., Strickland, T. C. T., Bosch, D. D., & Sullivan, D. G. (2007). Evaluation of SWAT manual calibration and input parameter sensitivity in the Little River watershed. *Transactions of the ASABE*, *50*(3), 843–855. http://doi.org/10.13031/2013.23149

Gouvêa, J. R. F., Valladares, G. S., Oshiro, O. T., & Mangabeira, J. a C. (2005). Comparação dos modelos digitais de elevação gerados com dados SRTM e cartas IBGE na escala 1:250.000 na região da bacia do Camanducaia no Estado de São Paulo. *Anais Do XII SBSR*, *7*, 2191–2193.

Gu, R. R., & Li, Y. (2002). River temperature sensitivity to hydraulic and meteorological parameters. *Journal of Environmental Management*, *66*, 43–56. http://doi.org/10.1006/jema.2002.0565

Helsel, D. R., & Hirsch, R. M. (1992). Statistical methods in water resources. *Elsevier*, *49*. Retrieved from http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0027065640&partnerID=40&rel=R5.5.0

Inchell, M. W., Rinivasan, R. S., & Uzio, M. D. I. L. (2013). SWAT user manual.

Ipcc. (2008). *Climate change and water: IPCC Technical Paper VI*. *Climate change and water* (Vol. 403). http://doi.org/10.1016/j.jmb.2010.08.039

Ipcc. (2013). Working Group I Contribution to the IPCC Fifth Assessment Report, Climate Change 2013: The Physical Science Basis. *Ipcc*, *AR5*(March 2013), 2014. Retrieved from http://ipccwg1.jp/AR4/meeting/pdf/SyR0407-siryo2.pdf\nhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:IPCC,+2014:+Climate+Cha

nge+2014:+Impacts,+Adaptation,+and+Vulnerability.+Part+A:+Global+and+Sectoral+Aspects.+Contribution+of+Working+Group

IPCC. (2014). Summary for Policy Makers. *Climate Change 2014: Impacts, Adaptation and Vulnerability - Contributions of the Working Group II to the Fifth Assessment Report*, 1–32. http://doi.org/10.1016/j.renene.2009.11.012

IPCC, I. P. C. C. (2007). *IPCC, 2007: Summary for Policymakers. in Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marqu eds. The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Retrieved from http://www.worldagroforestrycentre.org/sea/publications/searchpub.asp?publishid=1737\nhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:IPCC+2007:+Summary+for+policymakers#0

Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233–241. http://doi.org/10.1029/1998WR900018

Lenhart, T., Eckhardt, K., Fohrer, N., & Frede, H.-G. (2002). Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth, Parts A/B/C*, *27*(9-10), 645–654. http://doi.org/10.1016/S1474-7065(02)00049-9

Liersch, S. (2003). dew02.exe.

Lorz, C., Abbt-Braun, G., Bakker, F., Borges, P., B??rnick, H., Fortes, L., … Wummel, J. (2012). Challenges of an integrated water resource management for the Distrito Federal, Western Central Brazil: Climate, land-use and water resources. *Environmental Earth Sciences*, *65*(5), 1575–1586. http://doi.org/10.1007/s12665-011-1219-1

Merz, R., & Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of Hydrology*, *287*(1-4), 95–123. http://doi.org/10.1016/j.jhydrol.2003.09.028

Morawietz, M. (1997). User ' s Guide BFI. Oslo: University of Oslo.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900. http://doi.org/10.13031/2013.23153

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, *26*(7), 1465–1473. http://doi.org/10.1029/WR026i007p01465

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Srinivasan, R., & Williams, J. R. (2002). Soil and Water Assessment Tool User's Manual. *TWRI Report TR-192*, 412. Retrieved from http://swat.tamu.edu/media/1294/swatuserman.pdf

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., & Williams, J. R. (2009). Soil & Water Assessment Tool - Theoretical Documentation Version 2009.

Nikolakopoulos, K. G., Kamaratakis, E. K., & Chrysoulakis, N. (2006). SRTM vs ASTER elevation products. Comparison for two regions in Crete, Greece. *International Journal of Remote Sensing*, *27*(21), 4819–4838. http://doi.org/10.1080/01431160600835853

Oliveira, V., Makeschin, F., Sano, E., & Lorz, C. (2014). Physical and chemical analyses of bare soil sites in Western Central Brazil: a case study. *Environmental Earth Sciences*, *72*(2), 4863–4871. http://doi.org/10.1016/j.exppara.2012.04.011

Rodrigues, L. N. (2016). Personal correspondance. Brasília.

Rodrigues, L. N., & Torres, M. (2015). *Methodology to naturalized discharge in small watersheds with reservoirs and irrigation channels without discharge control,*. Brasília.

Shen, Z., Hong, Q., Yu, H., & Niu, J. (2010). Parameter uncertainty analysis of non-point source pollution from different land-use types. *The Science of the Total Environment*, *408*(8), 1971–1978. http://doi.org/10.1016/j.scitotenv.2009.12.007

Singh, V., Bankar, N., Salunkhe, S. S., Bera, A. K., & Sharma, J. R. (2013). Hydrological stream flow modelling on tungabhadra catchment: Parameterization and uncertainty analysis using SWAT CUP. *Current Science*,

104(9), 1187–1199.

Soito, J. L. D. S., & Freitas, M. A. V. (2011). Amazon and the expansion of hydropower in Brazil: Vulnerability, impacts and possibilities for adaptation to global climate change. *Renewable and Sustainable Energy Reviews*, *15*(6), 3165–3177. http://doi.org/10.1016/j.rser.2011.04.006

Spruill, C. A., Workman, S. R., & Taraba, J. L. (2000). Simulation of daily and monthly stream discharge from small watersheds using the SWAT model. *American Society of Agricultural Engineers*, *43*(6), 1431–1439. http://doi.org/10.13031/2013.3041

Tang, Q., Gao, H., Lu, H., & Lettenmaier, D. P. (2009). Remote sensing: hydrology. *Progress in Physical Geography*, *33*(4), 490–509. http://doi.org/10.1177/0309133309346650

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*. http://doi.org/10.1175/BAMS-D-11-00094.1

Terink, W., Lutz, a. F., Simons, G. W. H., Immerzeel, W. W., & Droogers, P. (2015). SPHY v2.0: Spatial Processes in HYdrology. *Geoscientific Model Development*, *8*(7), 2009–2034. http://doi.org/10.5194/gmd-8-2009-2015

Tolson, B. A., & Shoemaker, C. A. (2008). Efficient prediction uncertainty approximation in the calibration of environmental simulation models. *Water Resources Research*, *44*(4). http://doi.org/10.1029/2007WR005869

van Vliet, W. (2012). *Irrigation water strategies for catchment Buriti Vermelho : towards a higher water productivity*. Master thesis, University of Wageningen.

White, M. J., Harmel, R. D., Arnold, J. G., & Williams, J. R. (2014). SWAT Check: A Screening Tool to Assist Users in the Identification of Potential Model Application Problems. *Journal of Environmental Quality*, *43*(1), 208–14. http://doi.org/10.2134/jeq2012.0039

Xia, Y., Fabian, P., Stohl, A., & Winterhalter, M. (1999). Forest climatology: Estimation of missing values for Bavaria, Germany. *Agricultural and Forest Meteorology*, *96*(1-3), 131–144. http://doi.org/10.1016/S0168-1923(99)00056-8

Zhang, X., Srinivasan, R., & Bosch, D. (2009). Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. *Journal of Hydrology*, *374*(3-4), 307–317. http://doi.org/10.1016/j.jhydrol.2009.06.023

Appendix A

Detailed description of the data used during the thesis. Indicated are the data format, source and edits.

Data	Format	Description	Source	Edits
General data				
Boundary Brazil	Shapefile	Shapefile with the borders of Brazil to create the location map	IBGE - Brazilian Institute of Geography and Statistics	No edits, because this is no model input
Boundary Federal District	Shapefile	Shapefile with the borders of the Federal District	IBGE - Brazilian Institute of Geography and Statistics	No edits, because this is no model input
Location Brasília	Shapefile	Shapefile (point) with the location of Brasília	IBGE - Brazilian Institute of Geography and Statistics	No edits, because this is no model input
Hydrological data				
Buriti Vermelho Basin	Shapefile	Shapefile with the boundary of the Buriti Vermelho basin	ANA – National Agency of Water	Transfer of the coordinate system to wgs84 utm and from polygon shape file to grid so it can be used as mask within SWAT
Preto river basin	Shapefile	Shapefile with the boundary of the Buriti Vermelho basin	ANA – National Agency of Water	Transfer of the coordinate system to wgs84 utm and from polygon shape file to grid so it can be used as mask within SWAT
Preto river	Shapefile	Shapefile (line) with the location of the Preto river	ANA – National Agency of Water	Transfer of coordinate system to wgs84 utm.
Buriti Vermelho river	Shapefil	Shapefile (line) with the location of the Buriti Vermelho river	ANA – National Agency of Water	Transfer of coordinate system to wgs84 utm. Removal of small tributary streams in the streamfile.
Streams of Preto river	Shapefile	Shapefile (line) with tributataries of the Preto river	ANA – National Agency of Water	No edits, because this is no model input
Discharge data Preto river	Shapefile	Excel file of the stream gauge station located near Fazenda Limeira. Hourly average values (m^3/sec)	ANA – National Agency of Water	Transformation of format to attain vertical time series. Averaging hourly values to attain daily values.

Discharge data Buriti Vermelho river	Shapefile	Excel files of two stream gauge stations located located inside the Buriti Vermelho basin. Measured values per five minutes. Data logged as a change in discharge is significant(m3/sec)	Embrapa Agricultural research institute	Transformation of five minute data into a daily average in m3/sec.
Location stream gauge station	Shapefile	Shapefile (point) with locations of the stream gauge stations	Embrapa Agricultural research institute	Transformation coordinate value points to shapefile for the creation of a location map. Creating .txt file containing coordinates as input for SWAT
Meteorological Data				
Rainfall data	Excel files	Excel files of 103 rain gauges with daily rainfall values (in mm)	ANA – National Agency of Water	Selection rain gauges. (Time series must begin before 1984 and end in 2014; Max. 7% missing values). Transformation to SWAT input data (.txt-file per station)
Locations of rain gauges	Shapefile	Shapefile (point) with locations of the rain gauges	ANA – National Agency of Water	Coordinate points transformed into shapefile used for the creation of the rain gauge location map. Coordinates logged in .txt-file as input for SWAT.
Meteorological data	Excel files	Excel files with monthly values for averaged maximum and minimum temperature (°C), sun hours (hrs), relative humidity (%) and wind speed (m/sec).	INMET (National Institute of Meteorology)	Transformation from sun hours (hrs) to solar radiation (MJ/m2/day)as described in chapter 8.3
Locations meteorological stations	Shapefile	Shapefile (point) with locations of the meteorological stations	INMET (National Institute of Meteorology)	Coordinate points transformed into shapefile used for the creation of the rain gauge location map. Coordinates logged in .txt-file as input for SWAT.

Land-use and soil Data				
Land-use map	Shapefile	Shapefile with land-use of the years 1984 and 2014	Embrapa Agricultural research institute	Transfer coordinate system to wgs84 utm. Transformation shapefile to grid file. Lookup table created to translate land-use classes for SWAT.
Soil map	Shapefile	Shapefile with soil types and Excel files containing information about soil parameters	Embrapa Agricultural research institute	Transfer coordinate system to wgs84 utm. Transformation shapefile to grid file. Lookup table created to translate soil classes for SWAT.