Ungauged Catchment Hydrology: The case of Lake Tana Basin

B.Upul Janaka Perera

March, 2009

Ungauged Catchment Hydrology: The case of Lake Tana Basin

by

B.Upul Janaka Perera

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: Integrated Watershed Modelling and Management

Thesis Assessment Board

Dr. Ir. M.W.Lubczynski	(Chairman)	WREM Department, ITC, Enschede
Dr. Ir. J.Schellekens	(External examiner)	Deltaris, Delft
Dr. Ing.T.H.M.Rientjes	(First supervisor)	WREM Department, ITC, Enschede
Dr.A.S.M.Gieske	(Second supervisor)	WREM Department, ITC, Enschede

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

I dedicated this thesis to my father,
who is no more with us to see the completion of it
though he anticipated this to turn out.

Abstract

Estimating the water balance of a hydrologic basin often requires an effective procedure to quantify the contribution of the ungauged catchments. In this study a regionalizing based procedure is followed to quantify the flows of the ungauged catchments in the Lake Tana basin. Linking model parameters to physical catchment characteristics is a popular approach that enables the application of a conceptual model to an ungauged site. The HBV semi-distributed conceptual rainfall-runoff model is selected to simulate runoff of nine gauged catchments on a daily basis in the period 1994-2003. Some eight model parameters are selected for model calibration that also are used in the regionalisation procedure to simulate runoff from ungauged catchments. For nine gauged catchments the optimum parameter sets are derived through an automatic optimisation procedure based on Monte Carlo Simulation. By solving the multi-objective calibration problem that measures the different aspect of the hydrograph: (1) overall water balance (relative volume error), (2) overall shape of the hydrograph (Nash-Sutcliffe value). In the calibration procedure, for each catchment and for each single objective function best and worst values are used for rescaling to allow comparison. Thus, model performance is assessed through the use of rescaled objective functions that serve as criteria for selection of a best parameter set. In regionalisation, optimal model parameters are related to selected physical catchment characteristics (PCCs) that are used to estimate parameter values for ungauged catchments. Since PCCs from ungauged catchments generally differ from those of gauged catchments, also model parameter values will changed. By establishing the relations between values of HBV model parameters and PCCs, information was transferred from the gauged catchments to ungauged catchments. The transferred parameter sets were used to simulate the runoff from the ungauged catchments of the Lake Tana basin.

Four parameter regionalization methods (multiple regression, spatial proximity, similarity approach and sub-basin mean) were tested in this study to transfer model parameter values to the ungauged catchments. To evaluate the reliability of the simulations from the ungauged catchments, a water balance model of Lake Tana is developed at daily time step in the period 1994-2003. This model used area-volume and elevation-depth relations based on a bathymetric survey to simulate lake level fluctuations by calculation of the net inflow by estimating lake areal precipitation, lake areal evapotranspiration, inflow from gauged catchment, inflow from ungauged catchment and lake outflow. The water balance model of Lake Tana is evaluated by selected objective functions that are the Relative Volume Error and Nash-Sutcliffe objective functions as well.

Among nine gauged catchments in Lake Tana basin six of them are satisfactory calibrated under the criteria of relative volume error between ±5% and Nash-Sutcliffe value greater than 0.6 for the period of 1994-2000. Daily lake level simulation with inflow from ungauged catchment estimated from regression method shows the best performance with a relative volume error of -2.17% and a Nash-Sutcliffe coefficient of 0.92.

Key words: Regionalization, Water balance, HBV, Monte Carlo Simulation

Acknowledgements

First of all I thank the Load God for his endless mercy, grace and wisdom upon me during all these days in ITC and in all my life.

I expressed my sincere gratitude to my supervisor Dr.Irr. Tom Rientjes for his valuable support, guidance, encouragement and critical comments through out the research period. My sincere thanks also to second supervisor Dr. Ambro Gieske for his encouragement and comments to improve my research work and also former second supervisor Ing. Remco Dost for his assistance and cooperation until the last day in ITC.

I acknowledge the debt I owed to the National Water Supply and Drainage Board, for providing me this opportunity to continue my higher studies.

I wish also to thanks the Government of The Netherlands through the Netherlands Fellowship Program (NFP), which made this study possible by its financial support.

My appreciation is incomplete, if I do not mention the support and help extended by Lal Muthuwatta and Alemseged Tamiru Haile by means of constructive criticism, valuable suggestion and encouragements.

I would like to acknowledge Dr. Ben Maathuis, Ir. Gabrial Parodi and Ir.Arno Van Lieshout for their support during various part of the study.

I would also like to express my appreciation to all WREM department staffs and ITC community who helped me directly or indirectly during my study.

I would like to extend my gratitude to the Sri Lankan colleague for their laudable support throughout my stay in The Netherlands.

Above all, I wish to thanks my wife Treeni, my mother and mother-in law, for their moral support, without which I will never be able to complete this study. Apologies to my two daughters, Nathari and Nuveena. I wish you will forgive me when you realized the real reason behind parting you for such a long period. Dearest *thaththa* I wish you also forgive me not to being with you at you last hours.

Upul Janaka Perera
upuljanaka@gmail.com

Table of contents

List of figures

List of tables

1. Introduction

1.1. Background

The most important natural resource required for the survival of all living species is water. Water is limited, scarce, and also not well distributed in relation to the population needs. Hence proper management of water resources is essential to satisfy the current demands as well as to maintain sustainability. Water resources planning and management in the 21st century is becoming difficult due to the conflicting demands from various stakeholder groups, increasing population, rapid urbanization, climate change producing shifts in hydrologic cycles. Further variability in runoff in large ungauged international river basins affects water availability to downstream countries making management of transboundary water difficult and sometimes causing geopolitical tension. In most international river basin systems, the runoff variability of upper basins caused by natural or human interference is a serious concern in water resources management of all riparian countries. The upper Blue Nile basin is a typical example. The Blue Nile River Basin annually contributes about 60% of runoff to the Nile River (Conway, 2005) with its 10% areal occupancy. Reliable runoff information from this region is of great importance for sustainable management of water resources as the Ethiopian Highland is the major contributor to the Blue Nile River Basin.

1.2. Problem definition and importance of the study

A better understanding of the hydrological characteristics of different watersheds in the headwaters of the Nile River is of considerable importance because of the international interest in the utilization of its water resources, the need to improve and augment development and management activities of these resources, and the potential for negative impacts of climate change in the future. Conflicting views of water resource utilization and ownership has been challenging the development of appropriate management of the Nile River Basin for the countries that depend on its resources including Egypt, Sudan, and Ethiopia. Egypt's agricultural production and domestic requirements depends primarily (99%) on the waters of the Nile River and use of up to 55.5 billion cubic meters per year (Gheith and Sultan, 2002). Sudan's agriculture also relies heavily on these waters; while Ethiopia, at the Nile's headwaters, is interested in further developing the water resources to meet development needs and attain "water security" through equitable sharing among the Nile countries (Arsano and Tamrat, 2005). Only 5% of Ethiopia's surface water (0.6% of the Nile Basin's water resource) is currently being utilized by Ethiopia while cyclical droughts cause food shortages and intermittent famine (Arsano and Tamrat, 2005). There is an increasing demand for irrigation and hydropower development in Ethiopia to cope with the recurrent drought and its impacts and to increase agricultural production to cope with increasing population. Currently more than 90% of energy use in Ethiopia depends on the fuel wood (Kebede et al., 2006). This has resulted in extensive vegetation degradation which by itself effects the hydrological cycles.

Since there is no better infrastructure for water impounding systems or reliable irrigation schemes to compensate for prolonged low rain period, the Upper Blue Nile River Basin experiences drought and

famine. Further ungauged basins require planning of long term strategies such as hydropower generation, large-scale irrigation schemes and ecological protection as still many people largely depend on rain-fed agriculture and small-scale irrigation. Therefore estimation of runoff is important in gauged and ungauged catchment to understand the temporal and spatial variability of water yield for the local economies as well as the downstream countries. By scarcity of long term hydrological and meteorological data on the Upper Blue Nile basin, this study is concentrated only on the Lake Tana basin which is the source for the Blue Nile River. The basin occupies 10% of the Upper Blue Nile River basin and contributes 8% of the discharge (Kebede et al., 2006).

1.3. Scope of the research

Rainfall runoff models are often used to predict stream flow in space and time domain for operational and scientific investigations. By extrapolating and regionalization it enables us to simulate the catchment response of catchments for which time-series are not available (Heuvelmans et al., 2006; Seibert, 1999; Wagener and Wheater, 2006; Xu, 1999). Depending on several aspects the prediction of this discharge regimes in ungauged and gauged catchments bring along a given degree of uncertainty. In predicting gauged catchments several aspect can be considered: different model structures represent the real world differently and cause some degree of uncertainty, specific information and data required for these models cause uncertainty and the model parameters required for model calibration also contribute to the uncertainty issue in predicting stream flows. For better water management these predictions should be done accurately by reducing the uncertainties. Even though the prediction of discharge is possible and relatively simple in well gauged catchments, the prediction of discharge in ungauged catchment is complex and has higher degree of uncertainty. In gauged catchments the values of model parameters can be identified by calibrating the model against observed discharge. In ungauged catchments the observed data are not available or not sufficient for model calibration. Hence to predict the model parameters in ungauged catchments will depends on other sources of information.

The reliable estimation of continuous stream flow time-series in ungauged catchments has remained a largely unsolved problem so far (Wagener and Wheater, 2006), but significant insight has been gained in recent years. The establishment of the Prediction in Ungauged Basins initiative by the International Association of Hydrological Sciences (IAHS) shows that much has still to be done in this area (Wagener and Wheater, 2006). Several ways of obtaining the required information are addressed in literature. Analyzing the methods that have been attempted to estimate runoff from ungauged catchment from the previous studies, regionalization of model parameters, which is the process of transferring model parameter from comparable catchments to catchment of interest, was the common approach. It is possible to acquire necessary information in order to reduce model parameter uncertainty and uncertainty in predictions of discharge regimes in ungauged catchments.

The problem involved with developing even the simplest models lie primarily in data availability. Most established hydrological models are data intensive, yet the Lake Tana basin has limited rain gauge coverage, few long term temperature records, few gauged sub-catchments, and very scarce daily data. The size and complexity of the basin, together with the lack of data, is therefore a severe constraint to the application of sophisticated hydrological models. Application of a physically based model is practically not feasible because of their extreme data demand and instead conceptual models

are selected which have fewer parameters (Wale, 2008). Hence the HBV model is going to be adopted for this research with considering the following:

- The large number of application of this model, under various physiographic and climatological conditions, has shown that its structure is very robust and general, in spite of its relative simplicity (Lindström et al., 1997).
- The HBV model is a conceptual semi distributed model and a catchment can be partitioned into sub-basins, elevation zones and land cover types (Krysanova et al., 1999; Menzel and Bürger, 2002; SMHI, 2006). As the Lake Tana basin has mountainous topography this characteristic of HBV is more relevant to this study.
- It has been commonly proposed that parsimonious models are best suited to regionalization techniques (Crock and Norton, 2000), where HBV model has limited number of sensitive parameters and able to simulate the dominant hydrological processes.
- In the HBV model input data have been kept as simple as possible. This characteristic is suitable for this study area where data scarcity and data quality are the major problem.

From the Ministry of Water Resources (MoWR) and National Meteorological Agency (NMA), hydrological and meteorological data has been available for nine well gauged catchments and fifteen meteorological stations in Lake Tana basin respectively. By establishing relationships through the well gauged data between conceptual model parameters and physical catchment characteristics, parameters for the HBV model can be determined in poorly gauged catchments which will lead to a reduction of parameter uncertainty. Subsequently, this reduction of parameter uncertainty will lead to reduction of uncertainty in prediction of discharge regimes.

1.4. Previous studies and applications

A number of methods have been attempted to estimate the runoff from ungauged catchments during the past three decades. A summary of selected previous studies (see Table-list of previous studies of regionalization) shows that regionalization of model parameters was the most common approach. It is clearly indicated that each study used a conceptual hydrological model and there was no proper guide line for selection of physical catchment characteristics. The most common parameter regionalization method relates hydrologic model parameters with physical catchment characteristics (PCCs, hereafter) using regression methods. This approach seems effective in estimating the model parameters of ungauged basins with limited success (Heuvelmans et al., 2006; Kokkonen et al., 2003; Mwakalila, 2003; Yu and Yang, 2000). Several new methods were also attempted to better regionalize model parameters. These include the application of spatial proximity (Merz and Blöschl, 2004; Parajka et al., 2005), flow duration curves (Yu and Yang, 2000), basin similarity (Parajka et al., 2005), neural networks (Bastola et al., 2008; Heuvelmans et al., 2006), and regional calibration methods (Fernandez et al., 2000; Hundecha and Bárdossy, 2004). It is understood from the previous studies that the accuracy of estimation of runoff from the ungauged catchment cannot be guarantied by any particular method, but such depends on the quality of data available, specific basin responses, hydrological model and the regionalization method (Kim and Kaluarachchi, 2008).

1.5. Objectives

The general objective of this study is to apply a hydrological model and regionalization to simulate runoff from gauged and ungauged catchment. To achieve the above objectives, the following specific objectives were addressed.

- To make use of computer codes to incorporate automatic calibration in the HBV approach
- To apply Monte Carlo Simulation (MCS) for model calibration and parameter analysis
- To simulate the water balance of Lake Tana

1.6. Research questions

- What are the appropriate regionalization methods that can be used to extend the observed hydrologic regimes to ungauged basins?
- Which model parameters are most sensitive in MCS?
- Which physical catchment characteristics are available and suitable to relate to HBV model parameters?
- Can accuracy of inflows from ungauged catchments be defined?

1.7. Outline of the Thesis

This thesis contains seven chapters and is organized as follows:
Chapter 1 gives general introduction to the study with emphasis on water scarcity, problem and importance of the study and objectives of the study. Chapter 2 gives brief description of Lake Tana basin, data availability and the data quality. In chapter 3 insights is given in hydrological modelling by expounding the hydrological processes playing a role in rainfall runoff generation and describe the different aspect concerning the hydrological modelling. Further the approach of regionalization is described. Chapter 4 describe the methodology adopted in model calibration and regionalisation. In chapter 5 a detailed discussion is raised against the results of model calibration and regionalisation. Chapter 6 present the development of water balance model and simulation of lake water level. Chapter 7 finally discusses the conclusion and recommendations by this study.

2. Study area and data preparation

2.1. Study Area

The Lake Tana basin, one of the important sub basins in Upper Blue Nile river basin, located north western part of Ethiopia, occupies an area of 15,000 km^2 and drains to the Lake Tana which covers 3,060km^2. The basin lies approximately between latitude 10.75N and 13.0N and longitude 36.5E and 38.5E. Its altitude varies from 1784m at the Lake level to about 3400m of mountain in east of Lake Tana at the edge of the Ribb catchment. More than 40 rivers feed the lake. Among them are Gilgel-Abbay, Ribb, Gumera and Magetch that contribute more than 93% of the inflow. The only surface outflow of the basin is the Blue Nile River, which comprises 7% of the Blue Nile flow at the Ethio-Sudanese border (Conway, 2000).

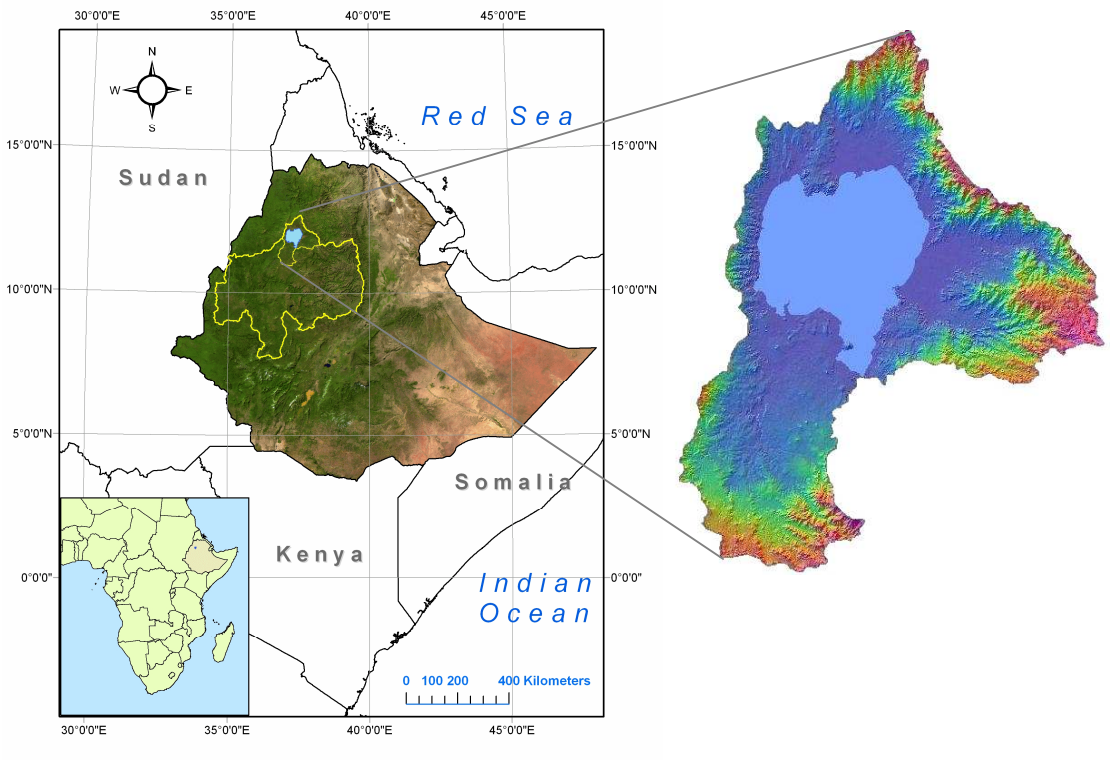


Figure 2-1: *Study Area*

The basin is characterised by Lake Tana which is located in a wide depression of the Ethiopian Plato and is surrounded by high hilly and mountainous terrain. The outflow of the basin is through a narrow valley at the south-east direction of the lake.

2.2. Climate

The climate of the region is 'tropical highland monsoon' with one rainy season between June and September. The seasonal distribution of rainfall is controlled by the northward and southward movement of the inter-tropical convergence zone (ITCZ). Moist air masses are driven from the Atlantic and Indian Oceans during summer (June - September). During the rest of the year the ITCZ shifts southwards and dry conditions persists in the region between October and May (Kebede et al., 2006). Even though it is associated with atmospheric circulation, the topography has also an effect on the local climate. The air temperature shows large diurnal variation but small seasonal changes and is comparatively uniform throughout the year.

Temperature, precipitation, wind and solar radiation are by far most important meteorological variable which drives the hydrological processes in a catchment. Precipitation, temperature and potential evapotranspiration are the main inputs for the HBV model considered in this research and described in Chapter 3. Great attention has to be paid in preparing and analysing the distribution of these data within the catchment in rainfall-runoff modelling practice. It is chosen to apply a lumped soil moisture routine in the HBV model since no detailed information about the catchments is available with respect to geographical location and elevation. Finally for every catchment average areal precipitation, temperature and potential evapotranspiration is derived as model input.

2.2.1. Precipitation

For this study daily precipitation was collected from the National Meteorological Agency (NMA) in Ethiopia for the period of 1994-2003 of fifteen rain gauge stations within and close to the study area is shown in Figure 2-2.

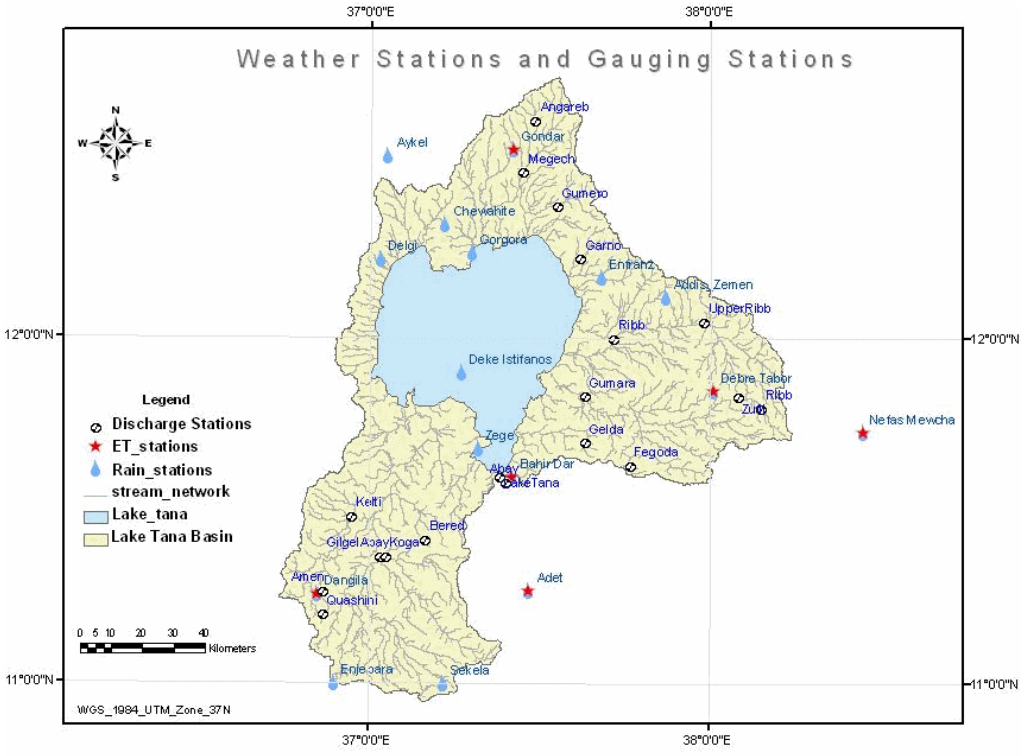


Figure 2-2: *Meteorological and gauging stations considered within and close to Lake Tana Basin*

There are missing and suspicious incorrect rainfall data for several stations and used most common methods described by Bras (1990) and Dingman (1994) such as normal ratio, weighted average and regression to fill the missing gaps and Double-mass curve analysis to analyse the consistency of suspected stations are used. The calculated annual average rainfall for each station from the above method is shown in Figure 2-3 with respect to their elevation. A detailed description about the gap filling techniques and determination of the areal rainfall over the catchment, accompanying issues and the used observation stations for averaging is given in Appendix B.

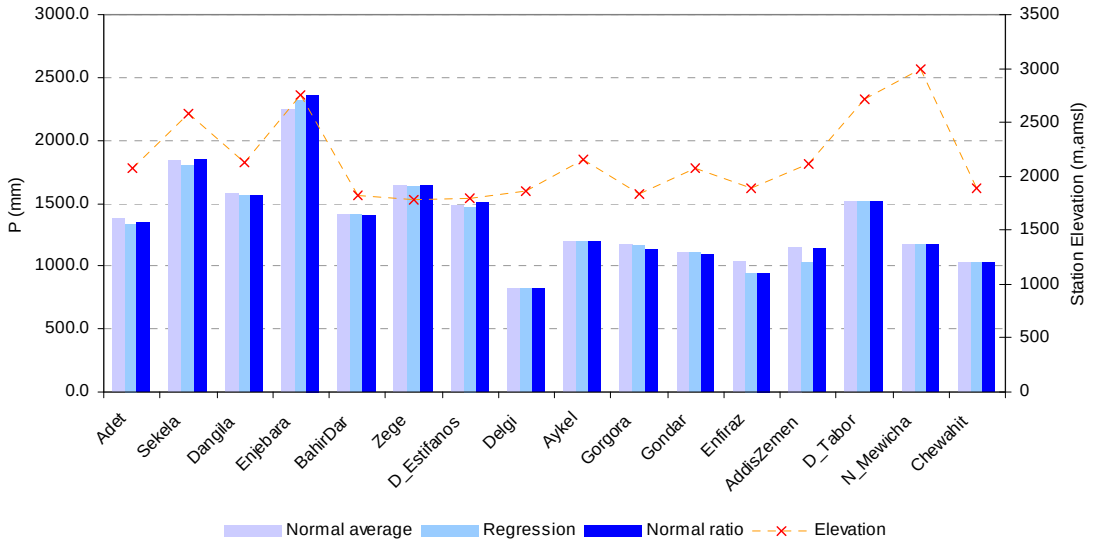


Figure 2-3: *Annual average rainfall in each meteorological station with its elevation*

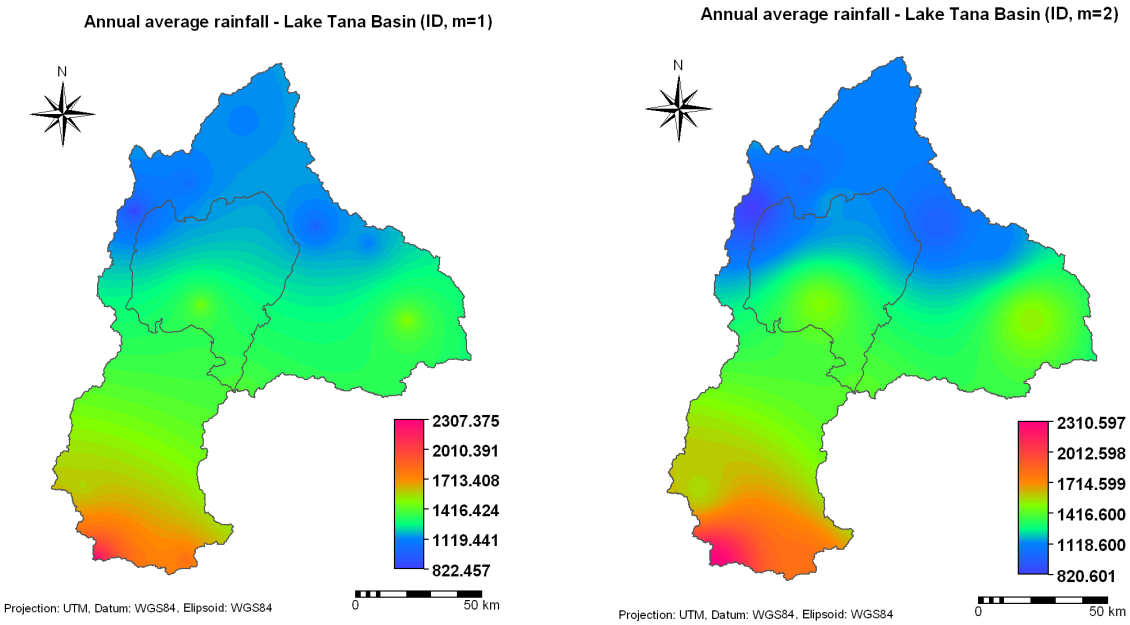


Figure 2-4: *Annual average rainfall distribution of Lake Tana basin*

These data series were used for calculating the areal rainfall over the catchments. To get insight in the variability of the precipitation across the catchments, the Annual Average Rainfall (mm/year) over the period 1994-2003 is shown in Figure 2-4.

The southern part of the Lake Tana basin is wetter than the western and the northern parts (see Figure 2-4). The rainfall-elevation relation over the entire basin show poor agreement. While examining the rainfall-elevation relationship on the southern part separately shows acceptable agreement and show great influence by the effect of elevation. But northern and eastern part of Lake Tana showed poor relation and can not explain such effect clearly. Figure 2-5 depicts this different of rainfall-elevation relation in two parts of the basin.

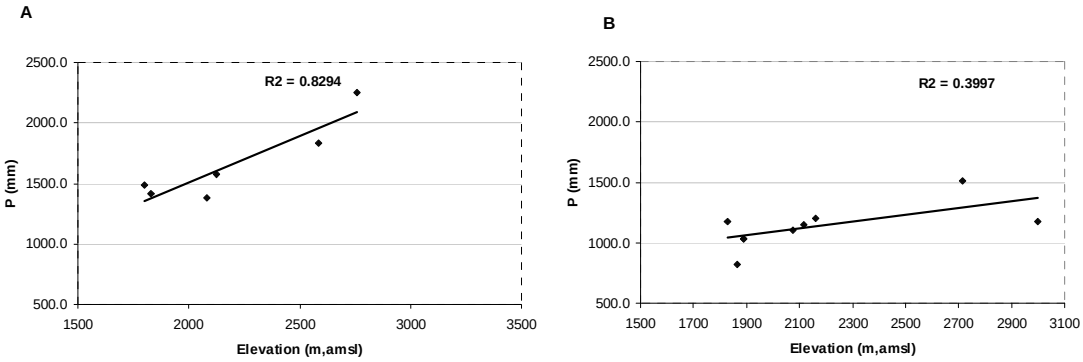


***Figure 2-5**: (A) Rainfall-elevation relation in southern part of Lake Tana. (B) Rainfall-elevation variation in northern and eastern part of Lake Tana*

Analysis of individual stations in the region shows that during the period from June to September it high rainfall is received and from October to May low rain. Figure 2-6 depicts the average daily rainfall distribution over the year for selected meteorological stations in the region.

2.2.2. Temperature

Daily maximum and minimum temperature data was collected from NMA and from the archives for the period of 1994-2003. Data of six meteorological stations within and close to the study area were chosen for analysis (see Figure 2-6).

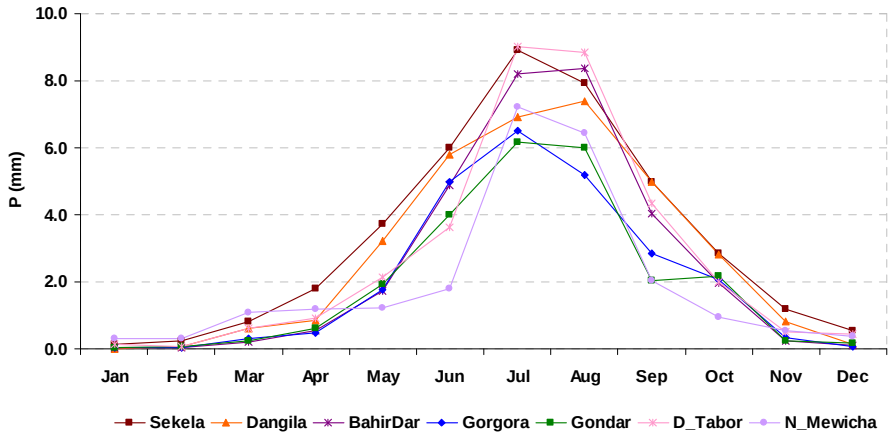


***Figure 2-6:** Average daily rainfall distribution throughout the year for several rainfall stations in Lake Tana basin*

The missing gaps of each data set were filled using the methods described in section 2.2.1. Areal temperature for each catchment was estimated by obtaining the weights by inverse distance interpolation. In order to get insight in the variability present across the catchment, the average daily temperature (ADT) is shown over the year (see Figure 2-7 [A]).

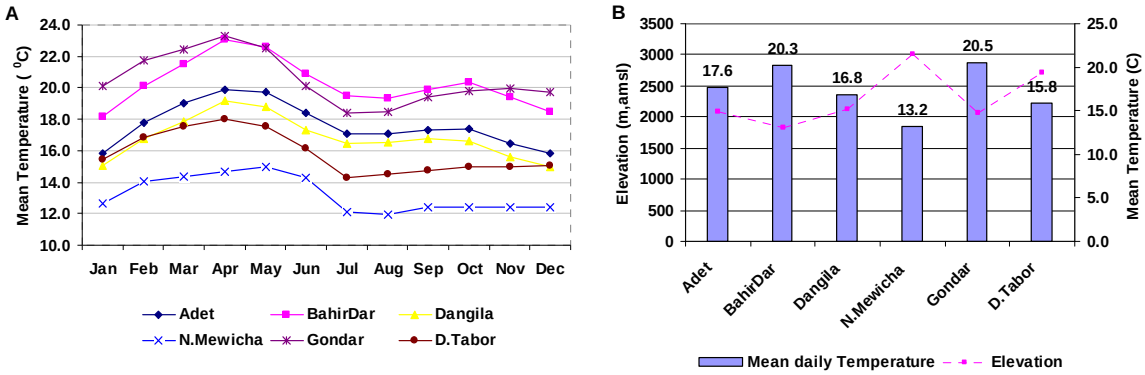


Figure 2-7: *(A) ADT variation for six different stations over the year during 1994-2003. (B) ADT variation of the stations with its elevation*

During the dry period ADT increases gradually from January to April and start to decline while the rainy season start. The minimum ADT for the respective station can be observed during July and August. During this period the basin received higher intensity of rain. Figure 2-7[B] clearly shows that the temperature of this basin is generally affected by the elevation: lower altitudes have high temperature and temperature declines while the altitude increases. The overall distribution of average temperature over the catchment shows lower temperature in southern parts.

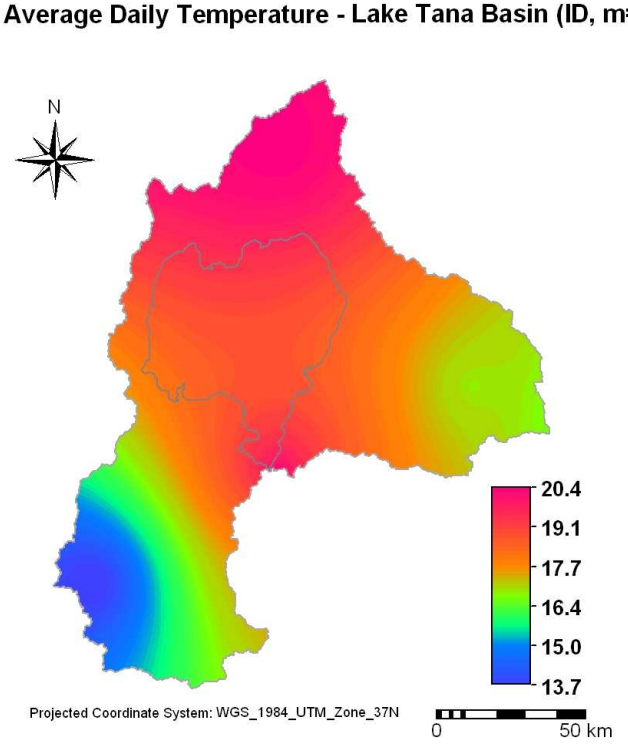


Figure 2-8: *ADT distribution over the Lake Tana basin*

2.2.3. Potential evapotranspiration

The last necessary input data for the model is potential evapotranspiration (PE). PE data is not found from the meteorological station or archives. Hence Penman-Montheith method was applied to estimate PE as this method is widely used in hydrological engineering application. The basic formula for calculating PE is shown in equation 2-1.

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad \textbf{[2-1]}$$

where
- ETo reference evapotranspiration [mm day^{-1}]
- Rn net radiation at the crop surface [MJ m^{-2} day^{-1}]
- G soil heat flux density [MJ m^{-2} day^{-1}]
- T mean daily air temperature at 2 m height [°C]
- u2 wind speed at 2 m height [m s^{-1}]
- es saturation vapour pressure [kPa]
- ea actual vapour pressure [kPa]
- es-ea saturation vapour pressure deficit [kPa]
- Δ slope vapour pressure curve [kPa °C^{-1}]
- γ psychometric constant [kPa °C^{-1}]

However for calculation PE, many variables are required such as: relative humidity, temperature, wind speed, altitude and sunshine hours. In many cases no observation stations are found with all data sets and if they were, most of the time the quality of the data is insufficient. Many data sets with missing time series and incorrect data were replaced with new values by the procedures described in Appendix B. Eventually, average areal PE is calculated for each catchment by finding the weight from each station using interpolation techniques. A detailed description about determination of the appropriate dataset of the four variables, accompanying issues and the used observation stations for averaging is also given in Appendix B.

In order to get insight in the variability across the catchment, the average annual evapotranspiration (AAE) is shown in Figure 2-9. The minimum average annual PE holds a value of 1291 mm whereas the maximum holds a value of 1448 mm. The distribution of evapotranspiration over the catchment shows low values in southern and eastern part and much higher in northern parts. By comparing with ADT, it shows significant effect for the fluctuation of evapotranspiration.

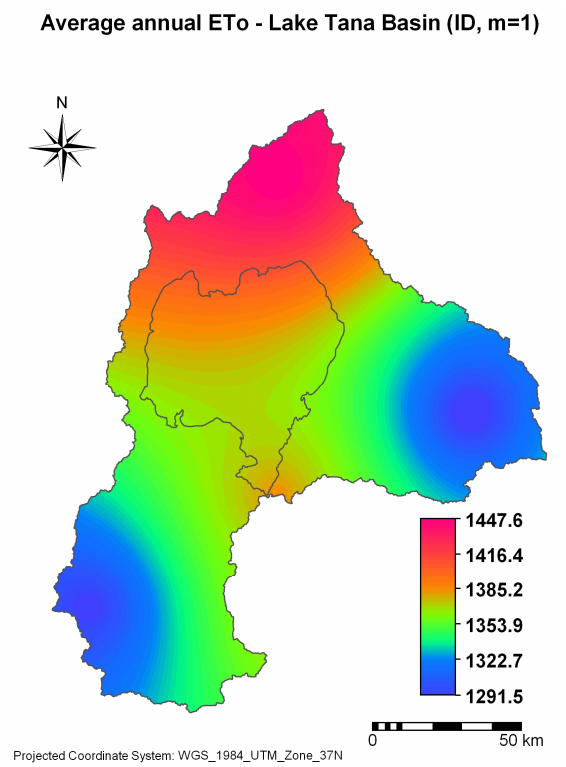


Figure 2-9: *Potential evapotranspiration distribution in Lake Tana basin*

2.2.4. Hydrological data

Many sub basins are gauged in Lake Tana basin by hydrometric stations to observe the runoff. The Lake level is monitored continuously with two level stations. The measuring devise of these stations are equipped with staff gauges and are recorded manually on daily basis. Some gauged catchments have several other gauging stations upstream that enable to analyse runoff correlations. Daily river discharge data series were collected over the catchment from the Ministry of Water Resources (MoWR) (see Figure 2-2).

According to the Tana Beles Sub-Basin Hydrological Study (SMEC, 2007) and Wale (2008) rating curves of many rivers in the Lake Tana basin are reliable (Gilgel Abbay, Koga, Megech and Kelti). However some of them are unreliable as the gauging stations are located middle of the flood plain (e.g. Ribb and Gumara rivers). One of another major problem is the sediment accumulation that disturbs the stage-discharge relation ship in many locations.

2.2.4.1. Analysing hydrological data

The collected river discharge data was screened to identify unreliable and spurious data. Screening was done graphically by comparison of hydrographs with sub basin hydrographs and double mass curve analysis with respect to rainfall. Some errors in river discharge data are described here.

In Figure 2-10A daily discharge data of Gelda River shows unexpected increase without any significant increase in rainfall The analysis of double mass curve clearly indicates the sudden

increment of the discharge during year 98/99, violating the well behaved relationship with respect to the rainfall of previous years (Figure 2-10B).

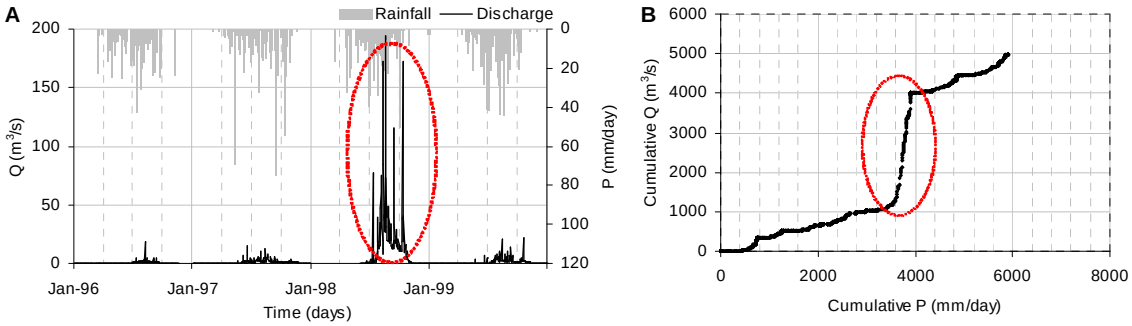


Figure 2-10: *(A) Unexpected river discharge in Gelda river 1998/1999, (B) Double mass curve for Gelda river with respect to rainfall for the same period.*

Another type of unrealistic river discharge data can be seen in Gilgel Abbay River during the recession period at the end of 1996. It shows sudden responses even though the nearby Koga river does not show such response. Further comparison with basin rainfall also does not show high rainfall events (Figure 2-11A).

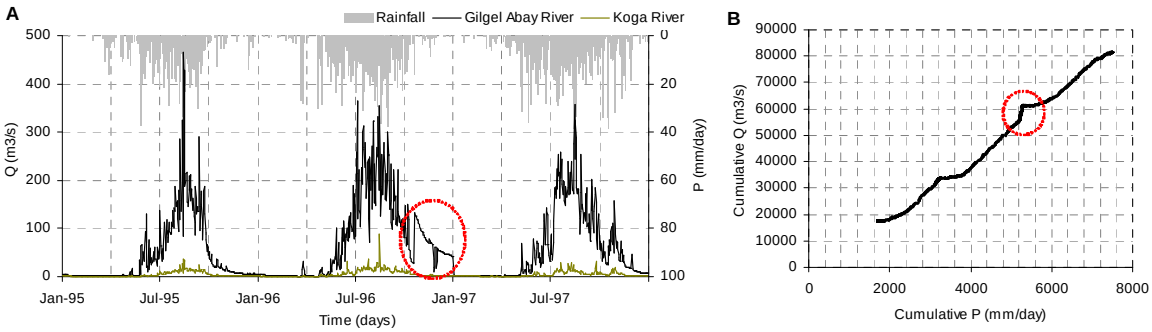


Figure 2-11: *(A) Suspicious River flow data in Gilgel Abbay River in 1996/1997, (B) Double mass curve of Gilgel Abey river with respect to rainfall for the same period*

In several years the peak river discharges of Rib River are trimmed and give constant value for a long period (Figure 2-12). This can be explained for that during extreme flooding the gauging station was submerged and caused difficulties in getting readings. or the river overflowed. Some of these errors are detected and explained also by Wale (2008)

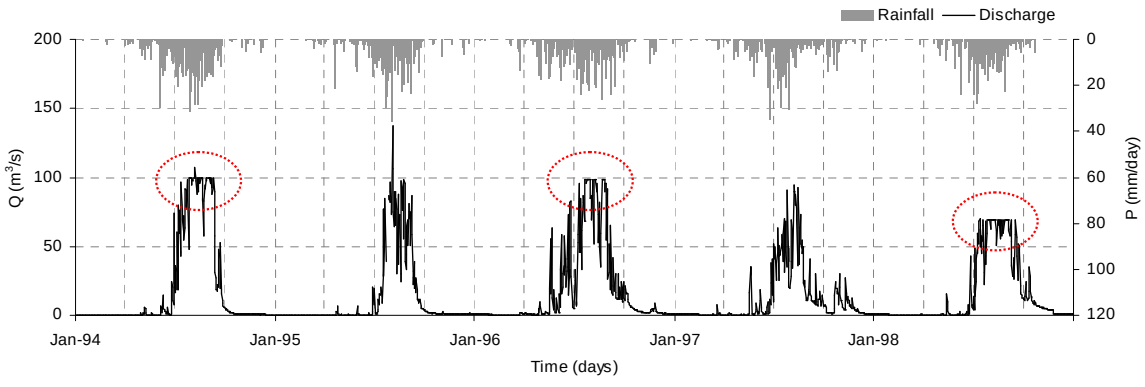


Figure 2-12: *Suspicious river flow data in Ribb river with respect to rainfall*

2.3. Delineation of gauged and ungauged catchments

A Digital Elevation Model (DEM) of 90m resolution from Shuttle Radar Topography Mission (SRTM-version 4) (http://srtm.csi.cgiar.org/) has been used to delineate the gauged and ungauged catchments by using hydro-processing tools in ILWIS and GIS software. A detailed procedure of delineation of catchment by ILWIS is available in Appendix C. The extracted 9 gauged catchments and 10 ungauged catchments are shown in Figure 2-13.

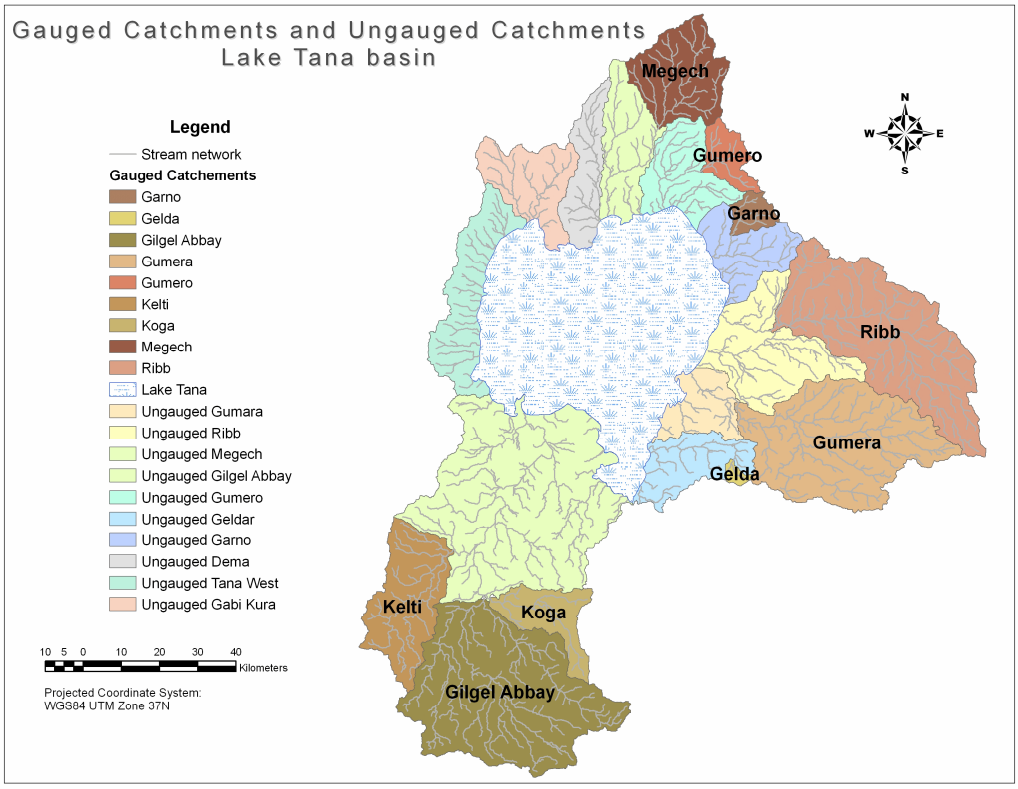


Figure 2-13: *Gauged and ungauged catchments in Lake Tana basin*

Results from catchment delineation, show that among the nine catchments seven catchments are partially gauged while catchments in the north-western part are completely ungauged. Nine catchments are selected as gauged catchments based on the availability of runoff data from 1994 to 2003. Representative areas of the gauged and ungauged catchment are in Table 2-1.

Table 2-1: *Major gauged and ungauged catchment delineated in the Lake Tana basin*

Gauged catchment		Ungauged catchment	
Catchment	Area (km^2)	Catchment	Area (km^2)
Ribb	1407.65	Ungauged Ribb	735.89
Gilgel Abbay	1657.41	Ungauged Gilgel Abbay	2071.95
Gumara	1281.31	Ungauged Gumara	286.63
Megech	531.18	Ungauged Megech	436.78
Koga	297.94	Ungauged Gumero	424.43
Gumero	163.32	Ungauged Garno	364.61
Garno	98.39	Ungauged Gelda	364.37
Gelda	26.13	Ungauged Dema	324.87
Kelti	607.64	Tana west	546.40
		Ungauged Gagi Kura	426.80

3. Hydrological modelling and regionalization

3.1. Hydrological modelling

3.1.1. Hydrological process

At first it is necessary to have insight in hydrological processes before hydrological modelling can be performed. The formation of the rainfall into stream flow involves many and complex hydrological processes in upstream area of the catchment. The basis of generating rainfall runoff processes can be found in the hydrological cycle, which is usually described in terms of precipitation, infiltration, evaporation, transpiration, surface runoff and groundwater flow. The precipitation over the catchment is the base of generating runoff. However after infiltration to the soil, the flow path becomes very unpredictable since the catchment runoff behaviour is closely related to the subsurface physiography, geometry and geology. Precipitation can be come out of the catchment as runoff through: overland flow, through flow and groundwater flow. For simulation of these processes a wide variety of hydrological models as well as applications have been developed over the past decades.

3.1.2. Classification of hydrological model

Many different types of hydrological models have been developed. Many of these models share structural similarities because of underlying assumptions, while some of the models are distinctly different. Therefore, these models are classified according to different criteria. Hydrologists traditionally proposed two kind of modelling approaches with their strong points and limitations: (1) physically based and (2) conceptual lumped models. Physically based models consist of formulations in terms of physical laws expressed in the form deterministic conservation equations for mass, momentum and energy. The equations are solved numerically by discretizing the hydrological system into smaller entities on a square or a polygonal mesh. However, accurately modelling of all processes of the hydrologic cycle becomes very complex, demands an eminent insight in hydrological behaviour and is very demanding for input data. Due to these properties it is a time-consuming and expensive. An example of such a model is SHE (Abbott et al., 1986).

As an alternative to physically based distributed models, conceptual lumped models are often used as robust prognostic tools at catchment scale. The model structures of these models are relatively simple and often are based on a series of interconnected reservoirs. Further these models are invaluable instruments for operational water management (e.g. reservoir operation, flood forecasting). The description of the reservoir's behaviour is kept simple in most structures and their responses are controlled by parameterizations that are rarely described in terms of physical principles such as gravity, piezometric heads or hydraulic conductivity and cannot be measured in the field. They must be estimated using a calibration procedure whereby the model parameters are fined tuned manually or automatically, by means of optimization algorithms, until the natural system output and the model output show an acceptable level of agreement. However, once environmental forcing conditions (e.g., switching from wet to dry conditions) or catchment characteristics (e.g., land cover pattern) change, the parameters usually need to be recalibrated (Reggiani and Rientjes, 2005). Because of the fact that

the required input and output data are usually easily available, consequently these models are mostly used in rainfall-runoff modelling. The HBV rainfall-runoff model is an example of a conceptual model (Bergström, 1995).

3.1.3. Model selection

The above model classifications must be considered when selecting the appropriate hydrological model in order to contribute to the stated objectives. At first sight, physically-based models are the most appropriate in modelling rainfall-runoff generation while it is the most elaborate way of modelling the rainfall-runoff processes. By the rapid improvement of computational power, the physically based models became practically applicable in the nineties. There has been considerable discussion regarding the pros and cons of this type of model by Beven (1989). In general it is concluded that physically based distributed hydrological modelling has clear limitations. Further, these models suffer from extreme data demand, scale-related problems (e.g. the measurement scales differ from the process and model scales) and over-parameterization. Therefore, in order to predict rainfall-runoff in gauged and ungauged catchments these models mostly are not practically applicable. Conceptual models normally perform at least as well as physically based models in predicting discharge regimes and furthermore they do not required huge complexity as physically based models do. Conceptual models on the contrary can be considered as a nice compromise between the need for simplicity on the one hand and the need for a firm physical basis on the other hand. This is due to the fact that these models usually are able to capture the dominating hydrological processes at the appropriate scale with accompanying formulations and therefore are very suitable when used in the process of regionalization. Furthermore, the HBV model has application in regionalisation (Hundecha and Bárdossy, 2004; Merz and Blöschl, 2004; Seibert, 1999) and it demonstrated to be effective. The improved HBV-96 model is used in this research is described in the next section.

3.2. HBV Model

The HBV model is a semi-distributed conceptual model and was originally developed at the Swedish Meteorological and Hydrological Institute (SMHI) (Bergström and Forsman, 1973). The general structure of HBV consists of three model components: (1) snow accumulation and snow melt, (2) the simulation of soil moisture and runoff, and (3) a response and river routing procedure. For the first two decades, only minor changes in the basic model structure were made. Experience has shown however, that the standard version of HBV had some major drawbacks which are outlined in Lindström et al. (1997). Therefore a re-evaluation has been carried out and a new model version has been developed. The HBV-96 model is the final result of this model revision (Lindström et al., 1997). Henceforward when “HBV model” is used, it is referred to the HBV-96 model.

The general water balance model is described as:

$$P - E - Q = \frac{d}{dt}[SP + SM + UZ + LZ + lakes] + ss \qquad \textbf{[3-1]}$$

Where

P - precipitation
E - evapotranspiration
Q - runoff
SP - snow pack

SM - soil moisture
UZ - upper zone
LZ - lower zone
Lakes - lake volume
ss – sink/source

In HBV model applications the water balance equation was simulated as follows. Precipitation over the catchment is calculated by weighting rain gauge measurements. The areal precipitation is then distributed over the elevation zones by correcting for altitude with a constant lapse rate. Evapotranspiration is computed as a function of the soil moisture conditions and potential evapotranspiration. When the soil moisture exceeds a storage threshold, water would evaporate at the potential rate. At lower soil moisture values a linear relation between the ratio of evapotranspiration and potential evapotranspiration and soil moisture is used. The general storage variable is formed by soil moisture storage and storage in the upper and lower response boxes. Recharge to groundwater is calculated through a non-linear relation between the ratio of recharge and precipitation and the soil moisture. The high flow of the catchment is described by the outflow from the upper nonlinear reservoir, while the base flow is governed from the lower response boxes which receive percolated water from the upper response box. Runoff from the catchment is given by the sum of the outflow from the two response boxes.

In this thesis a HBV model code has been developed based on the HBV-96 model version. The computer language FORTRAN is used and several reasons are on the basis for this decision. When using FORTRAN, adjustments to the model can be made which are not possible in the regular interface and such adjustment is the use of MCS for model calibration in this study. Booij (2005) formerly wrote the HBV model in FORTRAN for the Meuse basin and adjustments were made to the model based on the selected catchment, methods of routing routine, selected model parameters and model calibration method. A more profound description of the HBV-96 model is available in Appendix D.

3.3. Regionalization

Modelling of catchment runoff processes behaviour with rainfall-runoff models that merely are approximations of the processes taking place at the catchment scale. This is also the case for the HBV model. However, values for model parameters have to be established in order to simulate rainfall-runoff generation. Since for the HBV model it is not possible to directly determine the model parameter values, the parameters are normally estimated through a model calibration process by fitting the model output to observed discharge data. In many catchments observed discharge data is not available. Calibration of the model is therefore difficult and prediction of discharge regimes has a high degree of uncertainty. In order to allow prediction of discharge regimes in ungauged catchments, a method called regionalization is introduced. This method is used in this study to contribute to the objectives. There are many definitions of regionalization available in literature, but a generic definition as stated in Blöschl and Sivapalan (1995) is used most often. Regionalisation is the process of transferring information from comparable catchments to the catchment of interest.

3.3.1. Approach of regionalization

The transfer of catchment information in merely based on some sort of similarity between ungauged and gauged cachments. A number of methods have been applied to modelling ungauged basin such as spatial proximity (Merz and Blöschl, 2004), physical reasoning and statistical approaches (linking model parameters to PCCs). The spatial proximity method is based on the principle of catchments that are close to each other will likely have a similar runoff regime since climate and catchment conditions will often only vary smoothly in space. So the assumption is made that catchments are highly comparable with respect to topographic and climatic properties and therefore a particular model approach and with its parameters that calibrated for gauged catchments can be used to predict the discharge regime at the ungauged catchment. The most common techniques are nearest neighbour technique and kriging (Vandewiele and Elias, 1995). Researchers have concluded that geographical proximity does not guarantee hydrological similarity, and that methods based on physical reasoning are difficult to realize often due to the difference in scale at which the measurements are made and at which the model is applied (Bastola et al., 2008).

Another widely used approach to model ungauged catchments is to link MPs to PCCs. This method is referred to as a regional model method (Fernandez et al., 2000; Heuvelmans et al., 2006; Seibert, 1999; Wagener and Wheater, 2006). To broaden the modeling of ungauged catchment, many applications of conceptual rainfall runoff (CRR) models are based on functional relationships between MPs and PCCs. However such approaches are somehow elusive due to the existence of multiple optima and high interaction between subsets of fitted MPs (Kuczera and Mroczkowski, 1998). Some improved versions of conventional statistical approaches such as Weighted Regression, Sequential Regionalization and Regional calibration (Fernandez et al., 2000), have been discussed to overcome the problem of poor identification of MPs that mostly occur in over-parameterized CRR models. Wagener and Wheater (2006) discussed the uncertainties and problems of Sequential Regionalization and Weighted Regression methods for regionalization. However this approach considers the similarities of catchment characteristics and come up with three steps. First the selected model is calibrated for a reasonable number of catchments, where sufficiently long and informative observations of discharge regimes are available. Secondly, most commonly used regression equations are selected, which predict the model parameter values using one or a combination of PCCs. Each model parameter is based on a specific equation and thus, a specific regional model is obtained. Finally, these regional models will be used to estimate the parameter values of the ungauged catchment (Parajka et al., 2005).

3.3.2. Selection of regionalization method

The method based on special proximity can only be applied to ungauged catchments located close to a well gauged catchment. It requires a well gauged catchment in the near surrounding that often is considered as constrain. Kokkonen et al. (2003) mentioned that a simple transfer of parameters can outperform regionalization, even if catchments are hydrologically similar. In contrast to this approach, a regional model can be applied to various catchments even though it requires higher data demand than the spatial proximity approach. However it is easy to derive these data rather than finding long term discharge time series. There were two approaches addresses in literature with respect to the method of similarity of catchment characteristic. One approach predicts the discharge regime with a priori selected relationship between MPs and PCCs. The major disadvantage of this method is it's

difficult to identify the firm basis of such relationships in other studies to apply this study with sufficient confident. Further studies (Hundecha and Bárdossy, 2004; Merz and Blöschl, 2004; Seibert, 1999) rejected this method due to absence of satisfactory relationships. Hence in this study first the model was calibrated with respect to observed discharge and established the relationships between selected PCCs and MPs. Finally the model parameters for ungauged catchments were determined.

3.4. Model parameter selection

In order to determine the regional model, first model parameters have to be identified by calibrating the model against observed discharge. The major causes of difficulty in identification of model parameters are over parameterization and selection of parameters in the calibration. With respect to the HBV model, several modification have been made to the model structure to reduce the amount of parameters (Lindström et al., 1997), even though Merz and Blöschl (2004) mentioned reduction of over-parameterization is a critical issue. However many studies asses and conclude that HBV model is parsimonious enough. To establish relationships between PCCs all the parameters should not be used even though HBV model has more than 30 parameters. Such behaviour will induce extra effort in establishing statistical relationship. Therefore it is important to determine the most sensitive model parameters to be considered in the process of regionalization. In HBV model structure these processes are conceptualized by appointing appropriate model routines such as soil moisture routine which comprises the Horton overland flow and saturation overland flow, the quick runoff routing which comprises the macro pore flow and perched subsurface flow and the base flow routing which comprises the unsaturated subsurface flow and groundwater flow.

All parameters pertaining in a model routine do not affect to the same degree the rainfall-runoff transformation process. Therefore most sensitive parameters have to be identified in model calibration and subsequently in regionalization. By this analysis a chance is offered to developed effective relationships between parameters and PCCs. A number of studies used the HBV model approach and much experience was gained in demonstrating the most sensitive parameters. Hence, the parameter spaces from other studies were evaluated. Selected parameters and their spaces are listed in Table 3-1.

Table 3-1: *Selected model parameters and their priory range*

Name	Description and unit	Prior range	Default value
FC	Maximum soil moisture content [L]	100-800	Use a value for the region
BETA	Parameter in soil routine [-]	1-4	1
LP	Limit for potential evapotranspiration [LT^{-1}]	0.1-1	1
ALFA	Response box parameter [-]	0.1-3	0.9
KF	Recession coefficient upper zone [T^{-1}]	0.0005-0.15	0.09
KS	Recession coefficient lower zone [T^{-1}]	0.0005-0.15	0.01
PERC	Percolation from upper to lower [LT^{-1}]	0.1-2.5	0.5
CFLUX	Maximum value of capillary flow [LT^{-1}]	0.0005 – 2	0.5

3.5. Selection of physical catchment characteristics

The next challenge is to derive PCCs in order to establish statistically significant relationship between the model parameters and PCCs. A reasonable number of well gauged catchments with good quality

data of climate, hydrology, physiography, lithology and geology should be available to derive PCCs. However prior to selection of PCCs for regionalization, evaluation has to be done as there may be inter-dependency or inter-correlation between different PCCs. Therefore a preliminary list of PCCs should be composed based on the available data and the physical meaning of the model parameters. Then statistical analysis should be performed to identify the highly correlated PCCs. Correlation analysis and principle component analysis are the generally used methods to asses the correlations (Heuvelmans et al., 2006). Another more pragmatic approach is to evaluate those PCCs used in other studies and select them based on hydrological insight. Hence inter-correlation of the PCCs is also assessed inherently.

Several studies were evaluated in order to select appropriate PCCs. The detailed list of all the studies is shown in the Appendix A. Eventually the PCCs were evaluated and the selected ones are show in *Table 3-2* with the condition of availability of PCCs.

Table 3-2: *Selected PCCs*

Group	Parameter	PCC and Unit
Geography and physiography	*AREA*	Catchment area [km^2]
	LFP	Longest flow path [km]
	MDEM	DEM mean [m]
	HI	Hypsometric integral [-]
	AVGSLOPE	Average slope of catchment [%]
	SHAPE	Catchment shape [-]
	CI	Circularity index [-]
	EL	Elongation ratio [-]
	DD	Drainage Density[m/km2]
Land use	*CROPD*	Cultivated Dominantly[%]
	CROPM	Cultivated Moderately[%]
	GL	Grassland[%]
	URBAN	Urban[%]
	FOREST	Forest[%]
Geology and soil	*LEP*	Leptosol area [%]
	NIT	Nitosol area [%]
	VER	Vertisol area [%]
	LUV	Luvisol area [%]
Climate	*SAAR*	Standard annual average rainfall [mm]
	PWET	Mean precipitation wet season (Jun. to Sep.) [mm]
	PDRY	Mean precipitation dry season (Oct. to May) [mm]
	PET	Mean annual evapotranspiration [mm]

A brief description of each group of the PCC is given below.

1. **Geography and physiography**

Catchment area:

Area of the catchment is easily derived for each of the catchments. The amount of water reaching the outlet of the catchment mainly depends on its area. The model parameter *BETA* was reasonably

related to area according to the studies by Booij (2005) and Seibert(1999). In the study of Wale (2008) the area was related to *BETA*, *FC* and *LP* with high significance level.

Longest flow path:
The longest flow path is one of the outputs in catchment delineation processes in ILWIS. This indirectly is an indication of time for water to reach the gauging station. Wale (2008) established a reasonable relation between *FC* and longest flow path.

DEM mean:
Mean elevation is one of the frequently used PCC and is obtained as an output in the processes of delineation of catchment using SRTM DEM. Deckers (2006) mentioned that the elevation is highly correlated to slope. Thus, necessary attention should be made when establishing a relation between a certain PCC and elevation as well as slope.

Hypsometric integral:
This indicates the distribution of elevation across the catchment and simply calculated as:

$$HI = \frac{H_{mean} - H_{\min}}{H_{\max} - H_{\min}} \qquad \textbf{[3-2]}$$

where: H_{mean} – average altitude of the basin above mean sea level [m]
H_{max} – maximum altitude of the basin above mean sea level [m]
H_{min} – minimum altitude of the basin above mean sea level [m]

Average slope:
Slope is one dominant factor that controls the water flow velocity where a high slope result in high velocities that reduce the travel time of water to reach the catchment outlet. A percentage slope map was calculated using the function in ILWIS and categorised according to the FAO slope classes.

Class a: Level to undulation, dominant slope ranging from 0 to 8 percent
Class b: Rolling to hilly, dominant slope ranging from 8 to 30 percent
Class c: Steeply dissected to mountainous, dominant slope over 30 percent

Catchment shape:
Catchment shape is not a widely used PCC. This PCC is however expected to affect the hydrological processes at the catchment scale (Deckers, 2006). It is determined through formulae [3-3].

$$SHAPE = \frac{H_{\max} - H_{\min}}{\sqrt{AREA}} \qquad \textbf{[3-3]}$$

Circularity index:

The circularity index is calculated as the ratio of perimeter square to the catchment area.

$$CI = \frac{P^2}{A} \quad \textbf{[3-4]}$$

where P and A are perimeter [km] and area [km^2] of the catchment respectively.

Elongation ratio:

Elongation ration represent how the shape of the basin deviates from a circle. It is calculated by dividing the diameter of a circle with the same area as the catchment, by the length of the catchment.

$$EL = \frac{Dc}{L} \quad \textbf{[3-5]}$$

where *Dc*: is the diameter of the circle with the same area as the catchment, *L*: is the maximum length of the catchment along a line basically parallel to the main stem.

2. Land use

The characteristic land cover is one of the most used PCCs when establishing a regional model. This includes land cover type such as forest, grassland, crops and urban etc. It is a general concept that deforestation increases the soil erosion, since it changes the soil properties and infiltration rates. The land cover map, which has been updated by Landsat ETM+ and evaluated based on field data, used here was collected from ITC archives.

3. Geology and soil

Soil:

The soil map of the major soil groups of the catchment classified as per the FAO soil group was used for this study. This map was collected from the GIS department of EMWR. The dominant soils are described below.

Leptosols are extremely gravelly and stony. They can be seen over continuous rock in very shallow layers. These are azonal soils and common in mountainous regions,

Luvisols are soils that have higher clay content in the subsoil than in the topsoil as a result of pedogenetic processes (especially clay migration) leading to an argic subsoil horizon. These soils have high-active clays throughout the argic horizon. Further these soils have a medium to high storage capacity for water and nutrients and are well aerated.

Nitisols are deep, well-drained, red, tropical soils with diffuse horizon boundaries and a subsurface horizon with more than 30 percent clay and moderate to strong angular block structure elements that easily fall apart into characteristic shiny, polyhedric (nutty) elements. These soils have relatively high content of weathering minerals and surface soil may contain several percent of organic matter, in particularly under forest or tree crops.

Vertisols are churning, heavy clay soils with high proportion of swelling clays. These soils form deep wide cracks from the surface downward when they dry out.

All the above descriptions are taken from the FAO soil groups (WRB, 2007).

4. Climate

Standard annual average rainfall

The most commonly used characteristic is the standard annual average rainfall (SAAR) with respect to the climatic PCCs. For this characteristic the data are frequently available. However this is not commonly applied in other studies but in Deckers (2006) it proved to be a good indication for climatic variability.

Mean precipitation wet season and dry season

It is observed that there are two clear seasons for precipitation in the region, with high rainfall during June to September and low rainfall during October to May. Hence average daily rainfall in the wet season and the dry season was selected separately as climate PCCs.

Average annual evapotranspiration

The average annual evapotranspiration also has significant distribution over the catchment and varied from 1290 to 1450 mm/year. Kim (2008) has used this characteristics for the upper Blue Nile river basin and got reasonable relation with model parameters used in his study. In this study average annual evapotranspiration was used as a PCC.

4. Methodology

The approach of regionalization was applied to estimate the flow of ungauged catchments. First, the HBV model was calibrated against observed discharge to determine well performing parameter sets of gauged catchments. Next, a relationship was made between the model parameters (MPs) and physical catchment characteristics (PCCs) to establish the so called regional model that serves to estimate model parameters for ungauged catchments. Then the HBV model was used to simulate the discharged for ungauged catchments. Finally, the water level simulation model of Lake Tana was developed to increase our knowledge of the water balance of the Lake with the simulated ungauged catchments runoff. In the following subsection a description of the procedure is given.

4.1. Model Calibration

Any hydrological model must be proven for its reliability, accuracy and predictive ability. At the initial run the model probably will not give satisfactory result as the input data do not reflect the real world with enough accuracy. Hence every hydrological model requires adjustment of the model parameter values, hydrological influences and stresses in order to tune the model. The reliability of the model can be improved by fine-tuning termed ‘model calibration’ or simply ‘calibration’. The procedure of adjusting the model input parameters is necessary to match the model output with measured field data for the selected period and situation entered to the model (Rientjes, 2007).

4.2. Approach of calibration

The process of model calibration in order to identify the optimum model parameter set is done either by manually or by computer-based automatic procedures. In manual calibration, the user adjusts the parameters interactively in successive model simulations. As this approach mainly depends on the user’s experience, only intelligent steps will be made through the parameter space that will be an advantage. Some previous studies proved that this method demonstrated good model performance (Lidén and Harlin, 2000). However manual calibration is subjective and the parameter derived may be prone to be bias due to involvement of user’s experience. Also the process is very time-consuming and it does not have clear point at which the calibration process can be said to be completed. To overcome these disadvantages, automatic calibration has been pursued. In automatic calibration model parameters are adjusted automatically according to a specific search algorithm or optimization algorithm that considers numerical performance measures for the goodness of fit. Further, automatic calibration is fast with respect to manual calibration and the confidence of the model simulation can be explicitly stated (Madsen, 2000).

In this study calibration of the HBV model was done using Monte Carlo Simulation (MCS) to select a well performing optimum parameter set based on the objective functions, since the objective of this study was to establish a robust regional model rather than to optimally calibrate the catchment models. The MCS is a technique where through numerous model simulations with randomly generated model parameter sets, an optimum value for the objective function(s) is sought (Booij et al., 2007). For all

selected parameters a parameter space was determined through defining lower and upper boundary values (Table 3-1).

4.3. Objective functions

In model calibration the model parameters have to be adjusted until the observed natural system output and the simulated model output show an acceptable level of agreement. This goodness of fit is always evaluated through an objective function which is selected based on several criteria. These criteria should be selected properly to evaluate different aspects of the hydrograph. Some criteria that were considered in selecting the objective function in this study are as follows, see (Madsen, 2000):

1. A good agreement between the average simulated and observed catchment runoff volume (i.e. a good water balance)
2. A good overall agreement of the shape of the hydrograph

Based on these criteria relative volume error (RVE) and the Nash-Sutcliffe (NS) coefficient are used as objective functions to evaluate model performance.

Relative volume error

The relative volume error (RVE) criterion is defined as:

$$RVE = \left(\frac{\sum_{i=1}^{n} \left(Q_{sim,i} - Q_{obs,i} \right)}{\sum_{i=1}^{n} Q_{obs,i}} \right) \times 100\% \qquad \textbf{[4-1]}$$

where Q_{sim}: is the simulated flow, Q_{obs}: is the observed flow, i: is the time step, n: is the total number of simulation time steps of the calibration period. RVE ranges between $-\infty$ to $+\infty$ where with a zero value implies no difference between the simulated and observed discharge. The relative volume error between +5% and -5% indicate that the model performed well while relative volume error between +5% and +10% or between -5% and -10% indicate that model perform reasonably. However, at the same time the distribution of the discharge throughout the calibration period can be completely wrong. Therefore, this criterion should always be used in combination with another criterion that considers the overall shape agreement.

Nash-Sutcliffe coefficient

The Nash-Sutcliffe coefficient (NS) is used to evaluate the overall agreement of the shape of the simulated and observed hydrograph. NS measures the efficiency of the model by relating the goodness of fit of the simulated data to the variance of the measured data. NS coefficient is defined according to the following equation:

$$NS = 1 - \frac{\sum_{i=1}^{n}(Q_{sim,i} - Q_{obs,i})^2}{\sum_{i=1}^{n}(Q_{obs,i} - \overline{Q_{obs}})^2} \quad \textbf{[4-2]}$$

where Q_{sim}: is the simulated flow, Q_{obs}: is the observed flow, $\overline{Q_{obs}}$: is the mean of the observed flow, *i*: is the time step, *n*: is the total number of time steps used during calibration. NS can range between –∞ and 1and a value of 1 corresponds to a perfect match of model discharge to the observed discharge. The NS value goes to 0 indicate that the model predictions are as accurate as the mean of the observed data. When the value is less than 0 (between –∞ and 0) the observed mean is a better predictor than the model. This criterion is used frequently in various studies and when values range between 0.6 and 0.8, the model performed reasonably. The model is said to perform very good when the values is between 0.8 and 0.9 and it perform extremely well when the value is between 0.9 and 1 (Deckers, 2006).

4.4. Determination of optimum parameter set

The MCS requires a feasible parameter space to determine the values for the objective function while the HBV model simulates the discharge. Hence as discussed in section 3.4.1 sensible parameters and the feasible parameter spaces were determined for MCS in model calibration. In MCS there are huge numbers of parameter sets generated randomly and it is not know for which parameter values the model performs best. In order to evaluate which parameter set performs best over the two selected criteria, the value of each criterion was scaled over their own range determined through the calibration runs. The NS value was scaled based on its minimum and maximum value and scaling was done through equation [4-3].

$$C'_{NS,i,n} = \frac{C_{NS,i,n} - \min(C_{NS,i,ntot})}{\max(C_{NS,i,ntot}) - \min(C_{NS,i,ntot})} \quad \textbf{[4-3]}$$

where *C*: value for the criterion, *NS,i*: NS value for specific catchment, *n*: calibration run number, *ntot*: total calibration runs.

Since RVE varies between –∞ and +∞ and performance is best at a value of 0, positive values as well as negative values may occur. For RVE scaling was determined through equation [4-4].

$$C'_{RVE,i,n} = \frac{\left|C_{RVE,i,n}\right| - \max\left|C_{RVE,i,ntot}\right|}{\min\left|C_{RVE,i,ntot}\right| - \max\left|C_{RVE,i,ntot}\right|} \quad \textbf{[4-4]}$$

where *C*: value for the criterion, *RVE,i*: NS value for specific catchment, *n*: calibration run number, *ntot*: total calibration runs.

After having scaled NS values and RVE values, for each calibration run the lowest value among these two were selected. This is shown in equation [4-5].

$$C'_{i,n} = \min\{C'_{NS,i,n}, C'_{RVE,i,n}\} \quad \textbf{[4-5]}$$

where *C'*: scaled value of the criteria, *i*: specific catchment, *n*: calibration run number

Finally, the optimum parameter set for each catchment is determined by selecting the highest values of all selected minimum values as determined through equation [4-5]. The best performed parameter set was selected by equation [4-6] considering these two criteria.

$$C_i = \max\{\min(C'_{i,ntot})\} \quad \textbf{[4-6]}$$

where *C'*: scaled value of the criteria, *i*: specific catchment, *ntot*: total calibration run number.

After determine the feasible parameter range, it needs to determine the so called optimum run number. In MCS, at the optimum run number it is assumed that it's examine the entire range of parameter values. To determine the optimum run number the model was run for Gilgel Abbay catchment by increasing the run number starting from 500 runs up to 80,000. In the mean time the average of NS and |RVE| of the best 25 parameter sets was calculated in the decisive scale. The results are plot against the run number (see Figure 4-1).

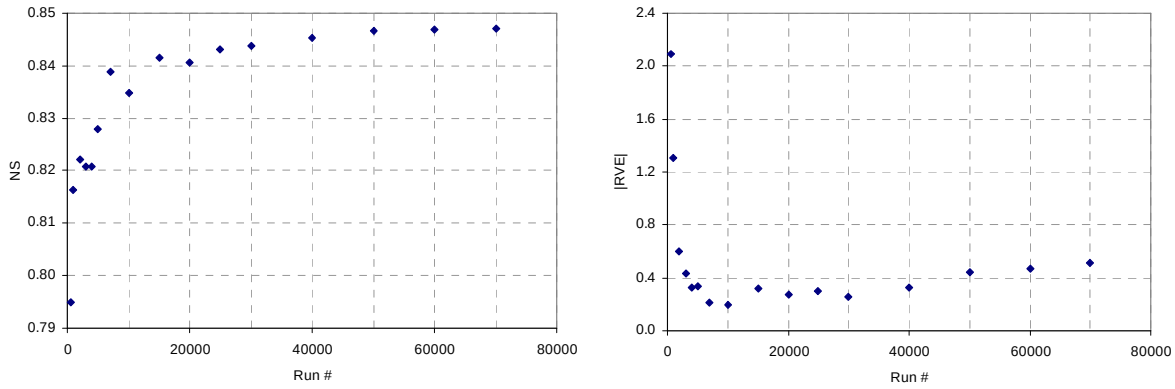


Figure 4-1: *Variation of average of best 25 NS and |RVE| value against run number*

The figure shows while increasing the run number the NS and |RVE| value come to a stable value. It can be concluded that the parameter space is well examined when the run number reach around 50,000 runs. Harlin and Kung (1992) also state that when the mean of the selected criteria shows stability implies the parameter space is well examined. Thus, for each calibration run number is fixed to 60,000 and selected the calibration period from 01-01-1994 to 31-12-2000.

4.5. Establishing the regional model

The relation between HBV model parameters and PCC allows us to understand and perhaps quantitatively predict how a change in physical properties of a catchment will affect its hydrological response (Mwakalila, 2003). In order to set up a regional model to predict the model parameters in ungauged catchments, a statistically significant and hydrologically meaning full relation should be established between PCCs and optimized MPs. After determining the MPs through model calibration

and selection of physical catchment characteristics as explained in section 3.4, a method for establishing the relationship is applied. The most commonly used approach is through regression analysis (Bastola et al., 2008; Booij et al., 2007; Heuvelmans et al., 2006; Kim and Kaluarachchi, 2008; Mwakalila, 2003; Parajka et al., 2005; Xu, 2003) and this approach is used in this study as well. A problem in this study is the limited number of available gauged catchments. In principle increasing the number of catchment will increase the reliability and the efficiency of the regional model. It is noted that in this study gauged catchments with RVE in between +5% and -5% and NS value greater than 0.6 were use to establish the regional model.

4.6. Regression analysis

A regression analysis is the application of a statistical procedure for determining a relationship between variables (Haan, 2002). In this procedure one variable is expressed as a function of another variable. The variable to be determined is termed the dependent variable while others are called the independent variables. In this study MPs were taken as dependent variable and PCCs were taken as independent variables. With respect to regionalization two type of regression methods are applied. One is simple regression method, which tries to determine a relationship between one independent variable and a dependent variable. The other type is called multiple regression method, which tries to assess multiple independent variables to explain the dependant variable.

Linear regression analysis

In this study linear regression is performed with respect to each model parameter. To optimize this simple relationship between MPs and PCCs the significance and strength of the relation was tested. First the correlation coefficient (r) was tested and in addition a t-test [eq. 4-7] was performed.

$$t_{cor} = \frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}} \qquad \textbf{[4-7]}$$

where, t_{cor}: t value of the correlation, r: correlation coefficient, n: sample size.

The following hypothesis was tested. The null hypothesis H_0 and the specific hypothesis H_1 are:
H_0: The correlation between the PCC and model parameters is zero, $\rho = 0$.
H_1: The correlation between the PCC and MP is not zero, $\rho \neq 0$
If $t_{cor} > t_{cr}$ the null-hypothesis is rejected (MP are associated with PCC in the population)

In order to determine the critical value t_{cr}, the degree of freedom df and α, a number between 0 and 1 to specify the confident level have to be determined. The default value for α is 0.05 for 95% confident interval. It must be note that the choice of level is largely subjective. In this study a significant level of $\alpha = 0.1$ was chosen to a two-tail test with n-2 degree of freedom. Using this information, for t_{cr} a value of 2.132 was found (critical value from t distribution table). In order to determine at what value of r the hypothesis is rejected the test statistic was solved. Result shows a value of 0.72, thus r greater than 0.72 or smaller than -0.72 results in a significant relationship.

Multiple regression analysis

The second method applied was based on the multiple regression analysis to optimize the linear relation with forward selection or backward removal method. Multiple regression is used to predict MPs from several independent PCCs. The forward entry approach the initially established regression model, which incorporates the significant best PCC will be extended by forcing a second independent variable in the regional model. This step will be accepted if the entry statistic (i.e. significance level, α) of both independent variables is not exceeded. The statistical tools are used to choose the most significant independent variable to be added. Further some more steps are executed, until the last added independent variable does not significantly contribute to the regression model. This equation is selected for further use.

The other method is backward elimination method that is applied in the event if all available PCCs exceed the entry statistic, i.e. at the first step no independent variable could be added when applying the forward entry method. This will terminate the regression analysis and no regression model would be established. But in the backward removal method all expected PCCs are forced into the model. Then the statistical tools will assess the removal statistic (i.e. significance level, α) and evaluate the eligibility to remove independent variables from the model. This will continue until the remaining independent variables do not exceed the exit statistics. In this way the multiple regression model is determined.

The r^2, the coefficient of determination, is not the only measure to rely on but also the significance of the relation has to be tested. In order to test the significance, the multiple regression equations can be tested by a test of significance of individual coefficients or by a test of overall significance.

4.7. Test of significance and strength

A hypothesis test has to be applied in order to determine if the regression equation is significant. It is possible to test several hypotheses. But for each test, assumptions have to be made and it is assumed that the error terms, ε, are not correlated and normally distributed. Further they have an average of zero and a constant variance. In this study these assumptions were made and two hypothesis tests were executed in order to test the significant of the regression equation. Those are null-hypothesis and specific hypothesis. Further the strength of the determined regression equation is also tested by r^2 called coefficient of determination. A detailed description of determination of significance of the regression equation with hypothesis test and the strength is described in Appendix E.

5. Results and discussion

5.1. Results of model calibration

Since it is difficult to estimate values of model parameters from the field measurements, all the model ranging from parsimonious lumped to complex distributed physically based need to be calibrated. In order to select an optimum parameter set, in this study MCS as calibration method was used. The HBV model was calibrated against the observed discharge by using the predefined parameter ranges as explained in section 3.4.1. However at the initial calibration, several catchments out of the total of nine catchments resulted in poor model performance with respect to the RVE. By screening the observed discharge data several unreliable and spurious data are identified for respective catchments. The methods described in Appendix B were adopted to correct those data points and corrected data has been used for calibration.

After calibrating the model for the period from 1994 to 2000 best performing parameter set is derived for each catchment with respect to the water balance and overall shape agreement of the observed discharge using RVE and NS respectively as objective functions (see equation [4-1] and [4-2]). Further by narrowing the parameter space some parameters are well identifiable. It can be observed that model is most sensitive in small ranges of parameter space for particular model parameters while model perform insufficient outside this range. To get an insight of well identifiable parameters a scatter plot of decisive scaled values (see equation [4-5]) against a priori selected parameter space of each MPs were made and Figure 5-1 shows selected eight parameters for catchment Gilgel Abbay. According to the results except for *ALFA*, others are not well identifiable since the model performed equally well for each parameter value within the a priori parameter space. An uneven distribution can be seen from the scatter plot of *FC*, thus it is also defined as poorly identifiable and the parameter space remained unchanged for further calibration. In the case of *ALFA* the scatter plot shows that that model performed well for lower values. Hence the parameter space of *ALFA* can be narrowed.

It is observed that for other gauged catchments the scatter plots show the same coherence and in order to assess all catchments at the same way the following approach is chosen to narrow the parameter spaces:

- For each catchment the best 6,000 out of 60,000 values of the decisive scaled values are derived. This for the reason that it is assumed that no good model performance will be realized when using a parameter set belonging to a poorer decisive scaled value.
- Hereafter, for every model parameter the respective parameter values are determined from which the minimum and maximum values of all 6,000 best runs are derived.
- These minimum and maximum values are used as a new boundary for the adjusted parameter space.

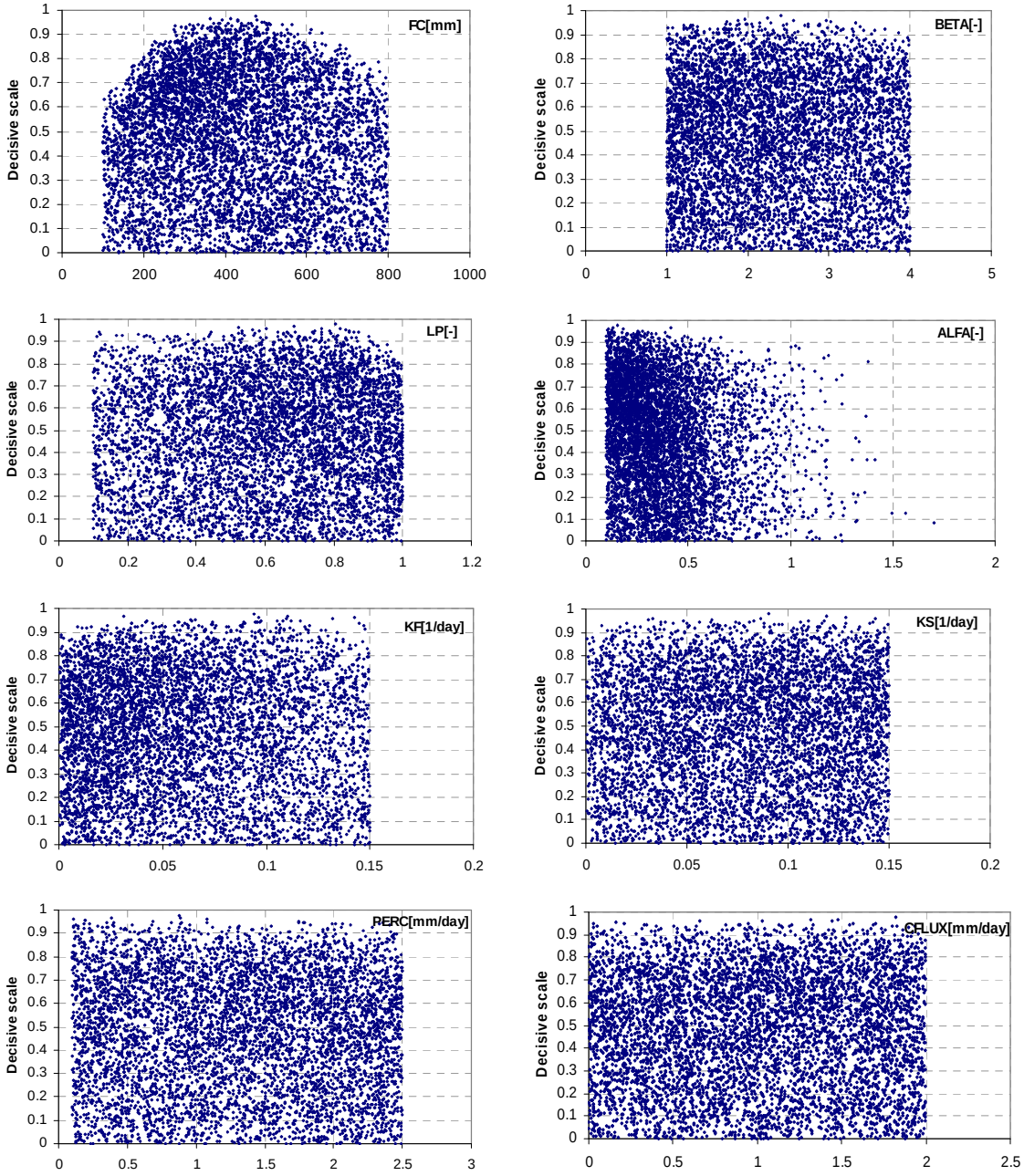
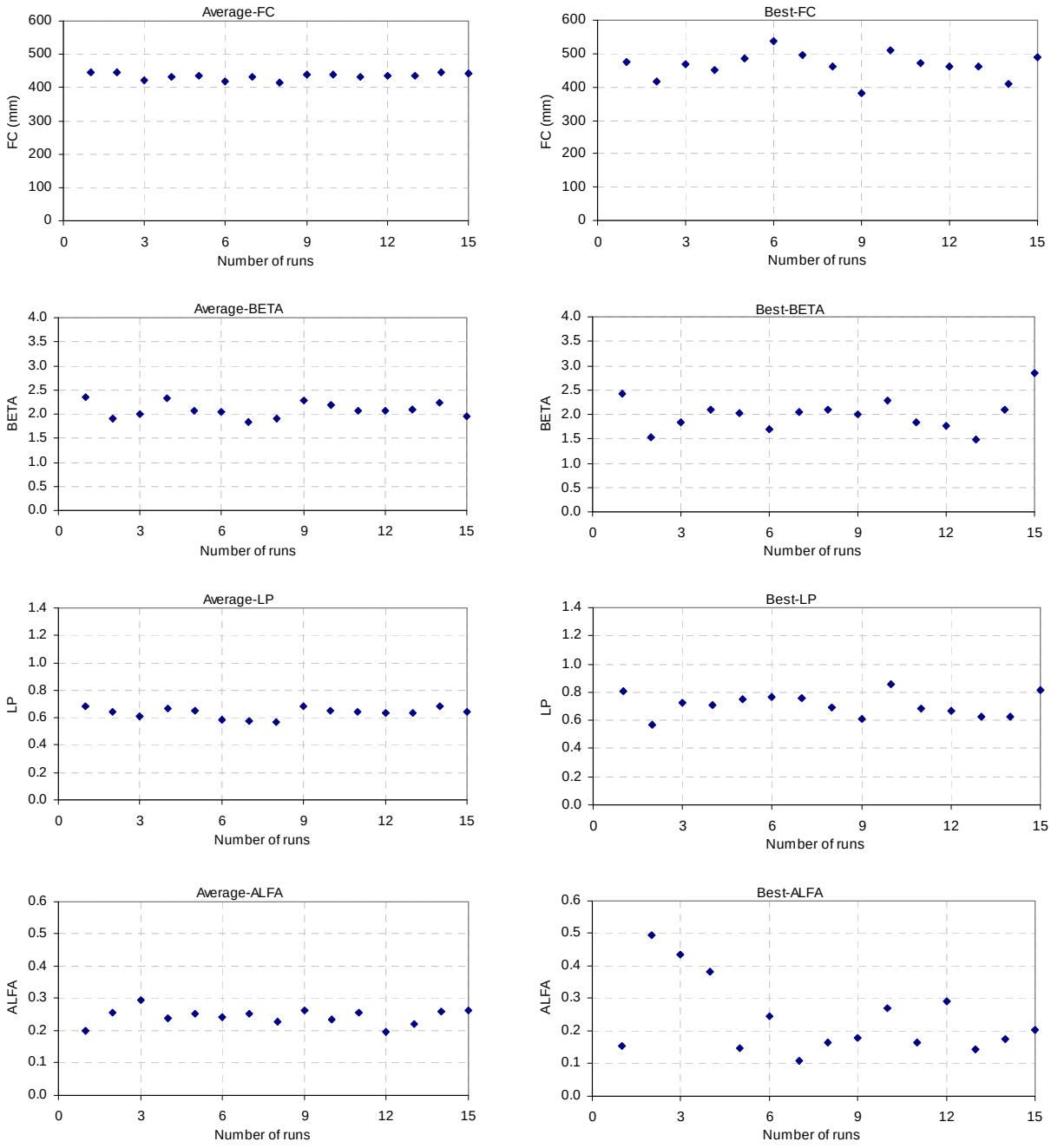


Figure 5-1: *Scatter plot of decisive scaled criteria against model parameters*

The above procedure is adopted for every catchment to narrow the parameter space. The posterior parameter spaces for each catchment are shown in Table G-2 in Appendix G. The result showed that if there is no major change in the parameter space, any change of parameter space will not help to find the best performing parameter set. The a priori parameter spaces were maintained for further use.

It showed that when the run number is 60,000, the entire parameter space was defined. But it can be observed that in different calibration run different well performed parameter sets can be found. To reduce this parameter uncertainty, the procedure was repeated 15 times and an optimum set for each run was defined. This was through a procedure of selecting and averaging over the 25 best performing

sets for each run. In the remaining of the procedure, to remove the outliers the variance over the 25 values were defined and parameter values that lies outside the standard deviation were denied. Graphical results of this procedure depicted in Figure 5-2 indicate that parameters are not unique but somehow converge to an optimal value. Finally the optimum parameter set was selected again taking the average of these 15 parameter values.



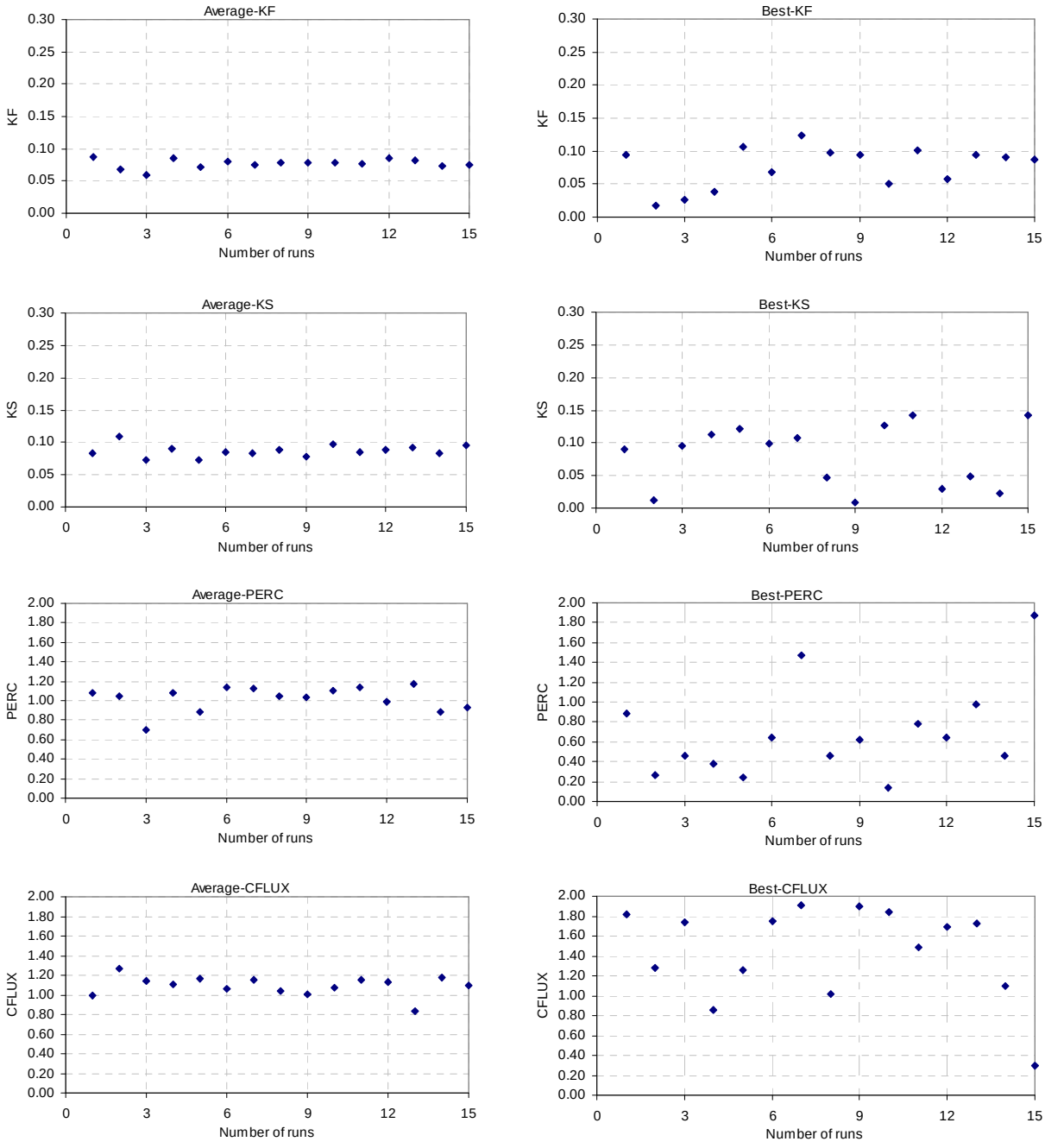


Figure 5-2: *Left hand side shows the average of best 25 parameters in each 15 calibration runs and right hand side shows best parameter value in each 15 calibration runs*

This procedure was applied for all the catchments and optimum parameter sets were established. Finally, the model was run and checked weather the NS and RVE values are in acceptable range. The model calibration results shown in Table 5-1 indicate that the model performance of Ribb, Gilgel Abbay, Gumara, Megech, Koga and Kelti catchments are satisfactory with RVE within ±5% and NS greater than 0.6. In this case for Ribb and Kelti catchments rainfall was corrected by 20% and 18% respectively.

Table 5-1: *Calibrated model parameters for gauged catchments (1994-2000)*

	Ribb	Gilgel Abbay	Gumara	Megech	Koga	Kelti	Gumero	Garno	Gelda
FC	309.03	434.39	349.86	193.04	730.05	196.62	469.23	221.25	141.14
BETA	1.23	2.08	1.31	1.56	1.34	1.60	1.10	2.58	1.20
LP	0.73	0.63	0.87	0.71	0.42	0.62	0.26	0.23	0.86
ALFA	0.31	0.24	0.25	0.29	0.41	0.28	1.08	0.27	0.51
KF	0.07	0.08	0.03	0.03	0.07	0.03	0.03	0.10	0.003
KS	0.10	0.09	0.07	0.09	0.05	0.10	0.13	0.11	0.15
PERC	1.09	1.02	1.44	1.47	1.63	1.53	2.32	1.61	1.41
CFLUX	0.60	1.09	0.72	0.79	0.74	0.83	0.39	1.35	1.00
NS	0.78	0.85	0.72	0.61	0.67	0.66	0.16	0.33	0.41
RVE%	-1.61	-0.35	-2.44	2.91	-0.06	-2.00	0.01	0.00	-0.06

The result of the calibration was not satisfactory for catchments with a relatively small area such as Gumero (163.32km^2), Garno (98.39km^2) and Gelda (26.13km^2). Hence these model parameters were ignored in establishing the regional model. Wale (2008) mentioned that the time of concentration, which is define as the length of time takes for water to travel from hydrologically most remote point to the outlet is relatively small and it is difficult to capture the quick runoff on daily time steps for these catchments. Time of concentration is measured according to the following equation and the basin time of concentration is shown in Table 5-2.

$$T_c = 0.7\left(\frac{L \cdot L_c}{\sqrt{S}}\right)^{0.38} \qquad \textbf{[5-1]}$$

where, T_c: is the time of concentration [hr], L_c: is the distance from the outlet to the centre of the catchment [km], L: is the length of the main stream [km], S: is the slope of the maximum flow distance path (Dingman, 2002).

Table 5-2: *Time of concentration for selected gauged catchments in Lake Tana basin (Wale, 2008)*

Catchment	Ribb	Gilgel Abbay	Gumara	Megech	Koga	Kelti	Gumero	Garno	Gelda
T_c [hr]	32.63	28.64	30.15	19.49	16.38	24.29	9.56	8.27	4.66

Further it was mentioned that those gauging stations are not placed at the catchment outlet, but at some location upstream that has easy road access. As such it is assumed that rainfall-runoff time series of those catchments can not be considered trustworthy. Figure 5-3 shows the model calibration results of accepted catchments.

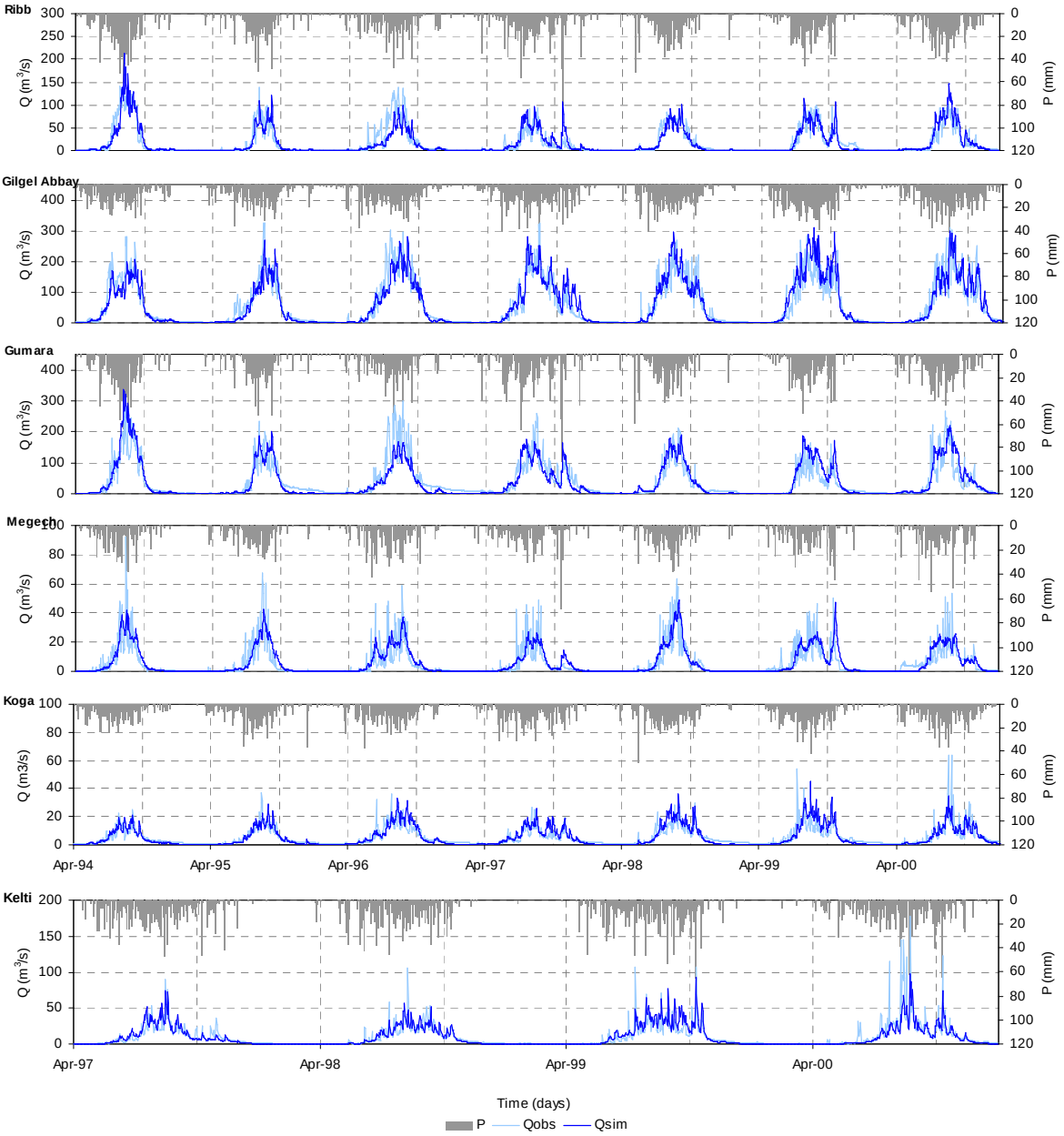


Figure 5-3: *Model calibration results of Ribb, Gilgel Abbay, Gumara, Megech, Koga (1994-2000) and Kelti (1997-2000) catchments*

5.2. Model validation

Representing the real world system by a model approach may not be accurate, since the real world is too complex. Models therefore are uncertain and models cannot be stated reliable when only one field situation is simulated. As such, it may occur that under different hydrologic stress conditions the model doesn't accurately represent the real world behaviors despite the fact that optimal and calibrated model parameters are used (Rientjes, 2007).

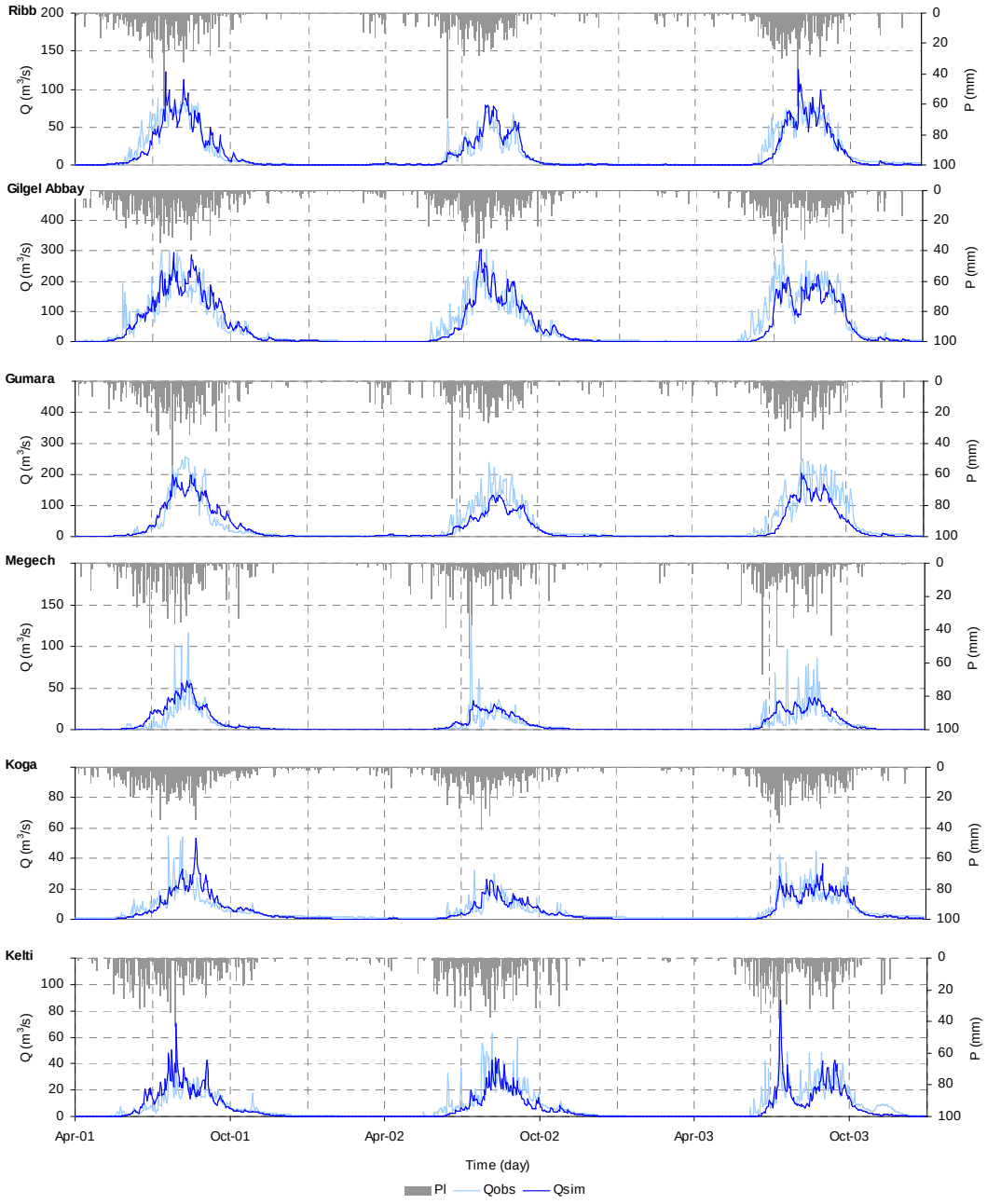


Figure 5-4: *Model validation results of Ribb, Gilgel Abbay, Gumara, Megech, Koga and Kelti catchment (2001-2003)*

The calibrated model is then run for the validation using the time series of meteorological variables that were not used for calibration. The model validation period runs from 2001 to 2003 (see Figure 5-4). Table 5-2 shows the model validation results from 2001 to 2003.

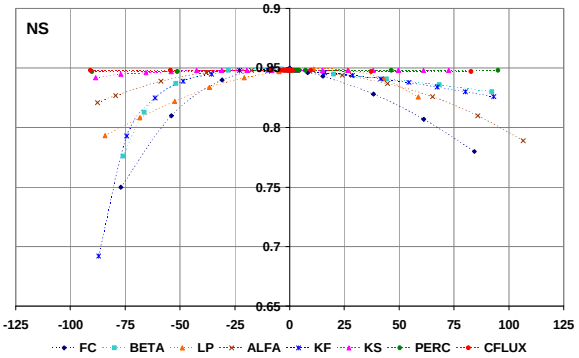
Table 5-3: *Model validation results from year 2001 to 2003*

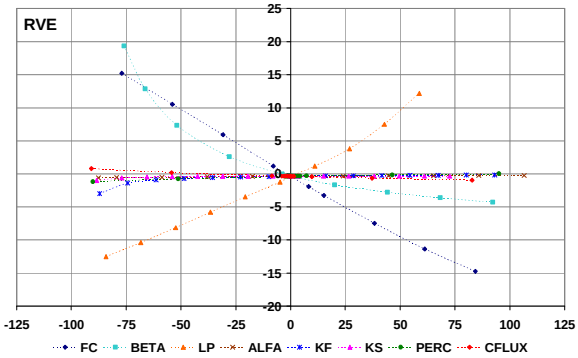
	Ribb	Gilgel Abbay	Gumara	Megech	Koga	Kelti
NS	0.87	0.85	0.79	0.51	0.65	0.67
RVE	3.55	-2.32	-9.87	2.87	-9.83	-5.30

5.3. Sensitivity of model parameters

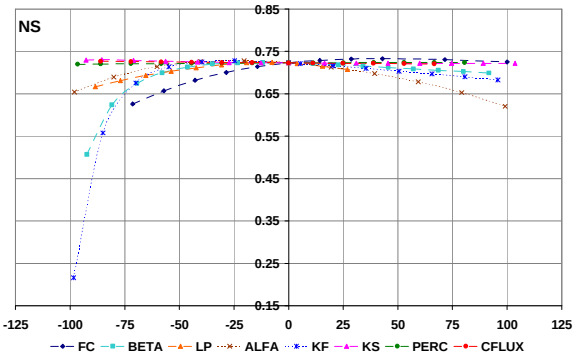
To get a clear understanding of the model behaviour with respect to the model outcome (i.e. model hydrograph) a sensitivity analysis is performed. Each of the eight model parameters contribute to conceptualizing the rainfall-runoff processes which all together result in simulating the hydrograph. Therefore if we change one of the model parameters the hydrograph will change. But every parameter does not contribute the equal amount of change to the hydrograph as described in section 3.4. When trying to establish relationships between model parameters and PCCs, it is most effective to investigate the most sensitive model parameters. Furthermore, it is useful to understand the influence of change in model parameter values on the hydrograph when evaluating the relationships. Figure 5-5 shows the sensitivity analysis for the Gilgel Abbay, Ribb, Gumara and Koga catchments.

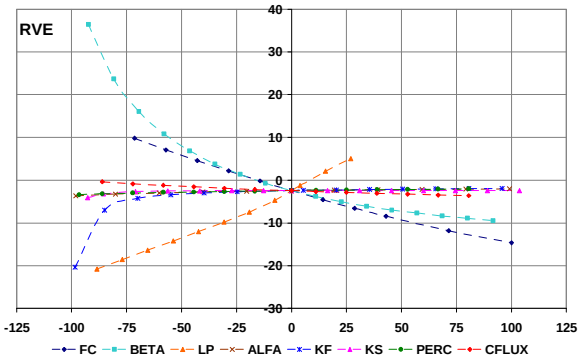
Gilgel Abbay catchment



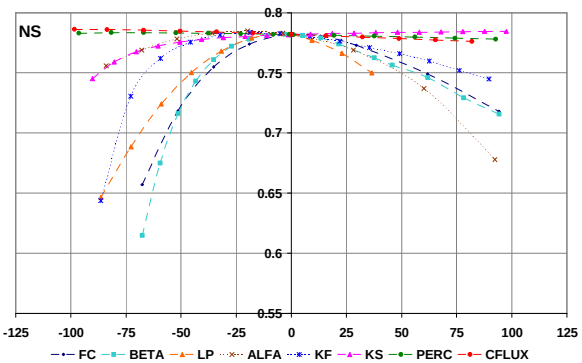


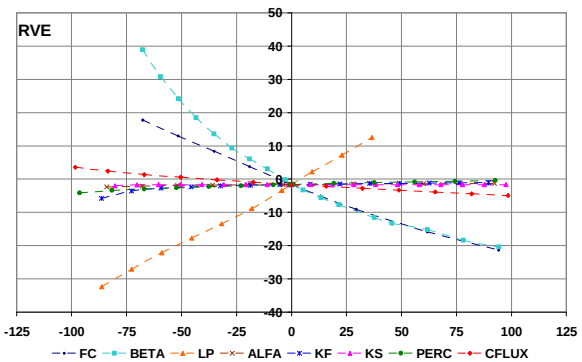
Gumara catchment





Ribb catchment





Kelti catchment

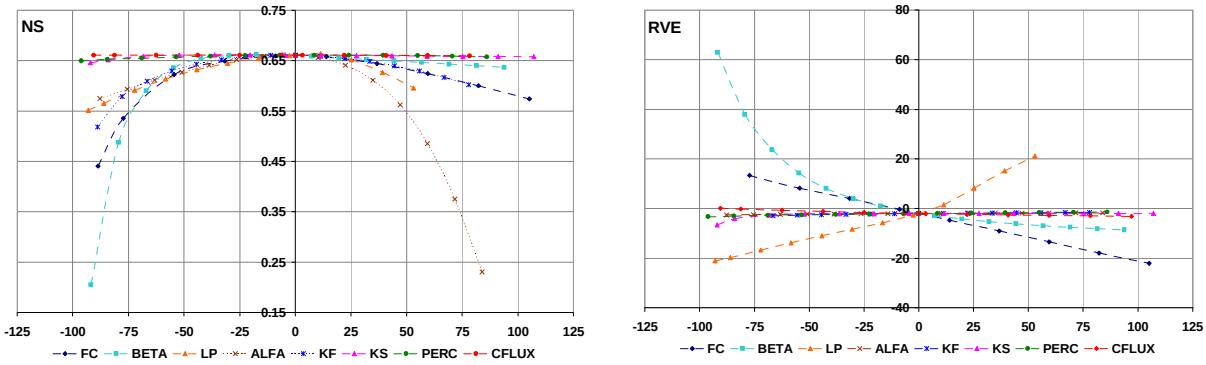


Figure 5-5: *Sensitivity analysis of Gilgel Abbay, Gumara, Ribb and Kelti catchments*

With respect to the NS value, among the 8 model parameters *FC*, *BETA* and *KF* act as the most sensitive parameter and show non-linear model performance to these parameters. *LP* and *ALFA* show relatively high sensitivity and also the non-linear behaviour to the model performance. *KS, PERC* and *CFLUX* act as less sensitive parameters and show less response to the model performance. The changes to *FC, BETA* and *LP* largely effect the volume of discharge and show higher variation in RVE, but not much effect from the other parameters. Further, the change of most of the parameter values resulted linear variation for RVE performance indicator.

Even though the relation between NS value with respect to the percentage deviation of parameter values show parabolic shape, it is not symmetrical. This can be explained by careful observation and visual interpretation of the hydrograph produced for each deviation of sensitive parameters (see Figures in Appendix J). For each model parameter an indication is given for which part of the hydrograph a specific model parameter is most sensitive.

It can be observed that small values of *FC* result in more responsive model behaviour. This is due to the reduction of storage capacity. This will result in a rapid drop of actual evaporation in dry season as water is quickly release from the system (see Figure J-1). Thus more water is released from the system and resulted in positive values for RVE. When the *FC* increases, the storage capacity will increase and the model increases the actual evaporation. Such causes delays in runoff response time and also reduces the discharge. This gives negative values to the RVE. Further the model does not behave the same way while the FC decreases or increases. When *FC* decreases, the model response very quickly and show significant changes on the hydrographs that results in large changes in NS and RVE values. But not such response is observed when *FC* increases.

Small values of *BETA* result in more infiltration from the soil moisture state to the quick runoff reservoir. In this case the potential to generate more discharge is higher (see Figure J-2). Hence it shows rapid variation of RVE and NS values while it reduces. But this higher response is not visible when increasing the value of *BETA*. It responses smoothly and the changes in NS and RVE are relatively small.

It is observed that lower values of *LP* can result in lower discharges; due to the fact that *LP* is a limit where above the evapotranspiration reaches its potential value. A high soil moisture storage the

evapotranspiration reaches its potential value and in total more precipitation will evaporate (see Figure J-3). This will give negative values for RVE. When the value increases, the actual evaporation reduces and the model releases more water to the system. But in this case also the reduction and increase of the *LP* value does not result in similar model response. It shows a different gradient for RVE when decreasing and increasing the *LP* value. In the decrement side of *LP* the gradient is low and in the increment side of the LP the gradient shows some higher value.

When changing the *ALFA* this affects the model performance with respect to NS value, but such is not clearly observed for RVE. When reducing the *ALFA* value, the model reduces the fast flow (see Figure J-4) and when *ALFA* increases it produces quick flow faster. This affects the shape of the hydrograph and also affects the NS value.

In changing KF, the NS value shows asymmetrical parabolic variation and RVE shows linear variation. Even though the variation of KF affects the NS value considerably, it does not affect much to the RVE. When reducing KF, the fast flow does not response well and also affects the hydrograph (see Figure J-5) and peaks are not well simulated. But when increasing the KF, it responses well and peaks are nicely matched. The model response differently when reducing and increasing the KF value and such causes the unsymmetrical curve for the NS values.

5.4. Effect of land cover changes

In order to assess the sensitivity of the response of the catchments to changes in the land use, different land use scenarios were generated and the model was run using the same meteorological input used in the calibration and validation periods while optimised parameters were keep unchanged. Two hypothetical land use scenarios were established. The first scenario was generated by reducing the forest percentage to 10% as uniformly distributed in each catchment to investigate the impact of deforestation on the runoff generation. The second scenario was used to investigate the effect of intensified afforestation and it was generated by assuming that the 90% of the catchment was forested.

By visual inspection of simulated hydrographs that clearly indicate that strong deforestation has an effect of increasing the peak flow to a considerable extent. (see Figure 5-6 and 5-7). Although the peak flow was increased consistently for all rainfall events, the extent to which it was increased was higher in the wet season than in the dry season.

Intensified forestation will increase the canopy interception and part of the precipitation will not contribute to runoff. Further it increases the transpiration, water holding capacity and increases delay of the quick runoff and overland flow. This will cause the reduction of high runoff.

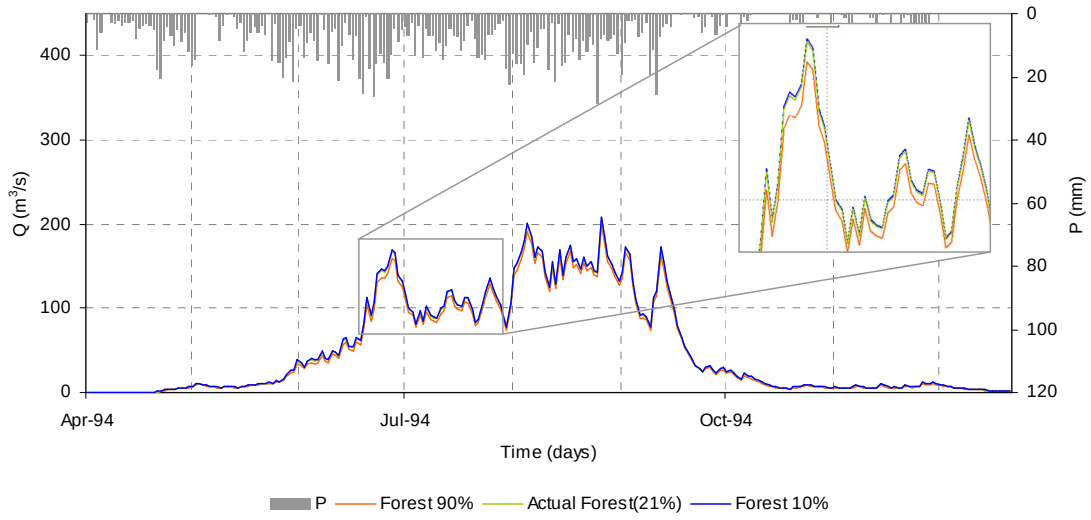


Figure 5-6: *Simulated hydrograph with different forest percentage in Gilgel Abbay catchment in 1994*

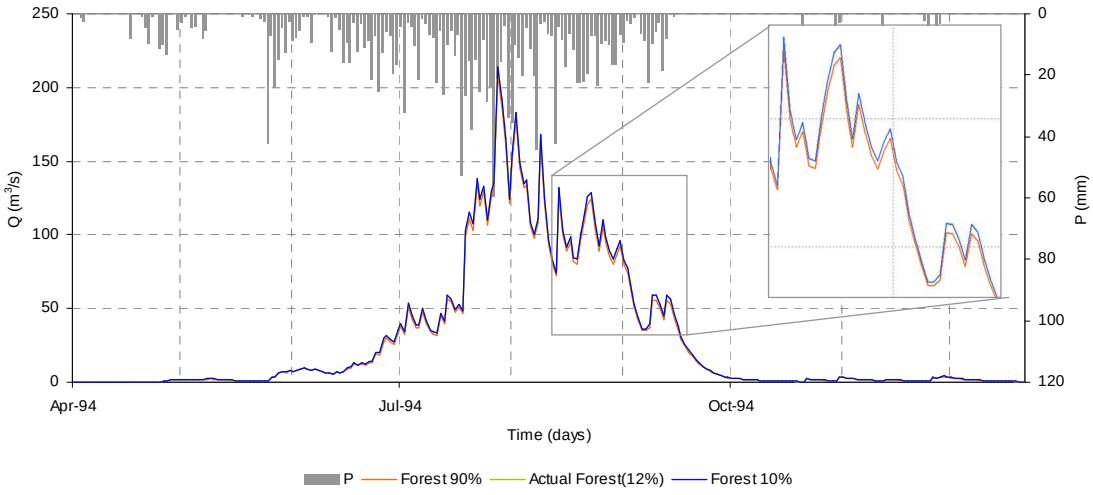


Figure 5-7: *Simulated hydrograph with different forest percentage in Ribb catchment in 1994*

5.5. Results regionalization

5.5.1. Simple linear regression

The correlation coefficient was established between PCCs and MPs in order to determine the significance of each relationship (Table). When the correlation coefficient lies outside critical values of -0.72 to 0.72, the corresponding correlation is significant. Thus, the null hypothesis is rejected. In Table 5-4 significant correlation coefficients are presented in bold text.

Table 5-4: *Correlation matrix between model parameters and PCCs for 6 selected catchments; significant correlation coefficients are in bold text*

	FC [mm]	BETA [-]	LP [-]	ALFA [-]	KF [1/d]	KS [1/d]	PERC [mm/d]	CFLUX [mm/d]
AREA	-0.18	-0.01	0.49	**-0.77**	0.13	0.62	**-0.82**	0.07
LFP	-0.17	-0.15	0.53	-0.51	0.25	0.52	**-0.74**	-0.16
MDEM	0.04	-0.58	0.33	-0.64	-0.11	0.66	-0.51	-0.51
HI	**-0.81**	-0.03	**0.77**	**-0.76**	-0.44	**0.77**	-0.44	-0.07
AVGSLOPE	-0.30	-0.48	0.39	**-0.75**	-0.31	**0.92**	-0.52	-0.45
SHAPE	0.65	**-0.90**	-0.42	0.12	-0.29	-0.04	0.43	**-0.83**
CI	**0.78**	-0.37	-0.55	0.58	0.37	-0.46	0.28	-0.37
EL	0.54	-0.59	-0.32	0.59	-0.01	-0.48	0.57	-0.66
DD	**-0.74**	0.23	0.64	-0.35	0.30	**0.77**	**-0.83**	0.06
CROPD	-0.39	**0.77**	-0.04	0.35	0.10	-0.36	0.20	0.69
CROPM	0.47	-0.71	0.03	-0.52	-0.22	0.27	-0.16	-0.55
GL	-0.18	-0.54	0.36	-0.19	0.08	0.61	-0.42	-0.67
URBAN	-0.53	-0.19	0.42	-0.71	**-0.72**	0.59	-0.05	-0.13
FOREST	0.67	-0.61	-0.60	0.47	0.20	-0.09	0.23	-0.63
LEP	-0.50	-0.36	0.20	-0.23	-0.59	0.57	0.14	-0.44
NIT	0.26	-0.31	-0.56	0.26	-0.49	-0.16	0.69	-0.26
VER	0.65	0.07	**-0.73**	**0.81**	0.18	**-0.88**	0.71	0.09
LUV	0.04	0.40	0.30	0.15	0.37	-0.55	-0.08	0.38
SAAR	0.45	0.52	-0.31	0.12	**0.75**	-0.29	-0.42	0.62
PWET	0.45	0.39	-0.21	0.07	**0.73**	-0.24	-0.46	0.49
PDRY	0.41	0.69	-0.45	0.20	0.71	-0.36	-0.31	**0.81**
PET	-0.13	-0.11	-0.23	-0.09	-0.59	0.13	0.43	-0.05

5.5.2. Multiple linear regression

It is assumed that the use of multiple PCCs will give better relation than the use of only one. Therefore relations between PCCs and MPs were assessed through multiple linear regression analysis. This was done by the forward entry method and backward removal method as described in section 4.6.

FC

FC[mm] corresponds to the maximum basin-wide water holding capacity of the soil. The value of *FC* can be estimated based on soil type and the rooting depth of the predominant vegetation and can further be refined in the calibration process (Hundecha and Bárdossy, 2004). The small *FC* values imply shallow hydrologically active soil depths, which may be realistic given that bare rock covers a substantial portion of the catchment areas in these regions (Merz and Blöschl, 2004). Zhang and Lindstrom (1997) showed that *FC* is correlated with climate, geological, geographical and hydrological conditions of the catchment. Merz and Blöschl (2004) revealed that *FC* has a positive correlation with porous aquifers and tends to increase the storage capacity of the catchment. Seibert (1999) found that there is a strong relation in between *FC* and lake percentage. Wale et al. (2008) found that *FC* has a negative relation with catchment area, longest flow path and average elevation of the catchment.

In this study *FC* showed significant positive correlation with *CI* and negative correlation with *HI* and *DD*. The highest correlation is with *HI*. The forward entry method was executed with *HI* as the initial variable. There was no other significant variable that improve the strength of the relation and the procedure was terminated. The regression equation was determined with only *HI* with R^2 of 66.3%. The statistical characteristics are shown in Table 5-5.

Table 5-5: *Statistical characteristics for the regression equation FC*

$FC = \beta_0 + \beta_1 \cdot HI$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	3520.82	0.0351	3.1317	1124.26	66.3%
β_1	-6651.21	0.0487	-2.8032	2372.70	

BETA

The HBV model covers a wide range of soil conditions in the runoff generation function with empirical parameters, *BETA* and earlier described *FC*. *BETA* describes how the runoff coefficient increases as the soil moisture approaches its limit of water holding capacity. *BETA* is thus more an index of heterogeneity than of soil properties in the basin. A *BETA*-value of zero implies that the basin is entirely lacking in water-holding capacity in the soil, whereas a high *BETA*-value indicates such homogeneous conditions that the whole basin may be regarded as buckets that overflow simultaneously when their field capacity is reached (Bergström and Graham, 1998).

The study done by Seibert (1999) showed that *BETA* and catchment area has positive correlation. Hundecha and Bárdossy (2004) showed that *BETA* correlated with soil type and land used. Merz and Blöschl (2004) found that *BETA* is negatively related to elevation and topographic slope. In the study of Deckers (2006) significant relation could not be established, and a backward elimination method was used to establish two regressions equations with Arable, URBAN and URBAN and SAAR . In the study of Wale (2008) *BETA* has significant positive correlation with catchment area.

In this study, BETA was negatively correlated to *SHAPE* and positively correlated to *CROPD*. As the highest relation is with *SHAPE*, the forward entry method was executed including *SHAPE* as the initial variable. The results of forward entry method showed that *BETA* is correlated with *SHAPE* and *HI* with R^2 of 96.02%. The statistical characteristics are shown in Table 5-6.

Table 5-6: *Statistical characteristics for the regression equation BETA*

$BETA = \beta_0 + \beta_1 \cdot SHAPE + \beta_2 \cdot HI$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	7.551	0.0100	5.85	1.2918	96.02%
β_1	-8.544	0.0429	-3.39	2.5233	
β_2	-0.036	0.0034	-8.50	0.0043	

LP

The next HBV model parameter in the soil moisture routine is *LP* that defines the minimum soil moisture at which the full potential evaporation takes place from the soil water. By a soil moisture below *LP* the actual evaporation reduces linearly to zero until the soil drains completely. In the study

by Hundecha and Bárdossy (2004) *LP* showed negative correlation with soil type. In the results of Wale (2008) *LP* has significant relation with catchment area. In this study LP has significant positive correlation with *HI* and negative correlation with *VER*. The forward entry method was executed with highly correlated *HI* as the initial variable. This result showed that *LP* was correlated to *HI* and *LUV* with R^2 of 91.1%. The statistical characteristics are shown in Table 5-7.

Table 5-7: *Statistical characteristics for the regression equation LP*

$LP = \beta_0 + \beta_1 \cdot HI + \beta_2 \cdot LUV$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	-2.2435	0.0258	-4.13	0.5432	91.1%
β_1	5.8697	0.0133	5.27	1.1141	
β_2	0.0027	0.0471	3.26	0.0008	

ALFA

In the response routine of HBV model, *ALFA* is the measure of the non-linearity in the upper reservoir. Booij (2005) indicated that small catchments with steep hills and low permeable soils generally result in more non-linear behavior in the fast flow mechanisms than large sub-basins with flat terrain and high permeable soils. Hundecha and Bárdossy (2004) established a positive relation with soil type and land use. But in the study of Wale (2008) and Deckers (2006) any significant relation with *ALFA* was not found and backward elimination was performed. Wale (2008) established two regression equations with bare and hilly percentage and logarithm of average altitude and bare land percentage. Deckers (2006) established a relation with elevation, hypsometric integral and permeability percentage. In this study *ALFA* has significant positive correlation with *VER* and negative correlation with *AREA*, *HI* and AVGSLOPE. Therefore optimisation of the linear relation with forward entry method was executed initiating with variable *AREA*. After adding the catchment characteristic *URBAN* the R^2 increased up to 95.1% and this regression equation was accepted. The statistical characteristics are shown in Table 5-8.

Table 5-8: *Statistical characteristics for the regression equation ALFA*

$ALFA = \beta_0 + \beta_1 \cdot AREA + \beta_2 \cdot URBAN$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	0.45233	0.0003	18.63	0.02428	95.1%
β_1	-0.00009	0.0251	-4.17	0.00002	
β_2	-0.73650	0.0341	-3.71	0.19865	

KF

KF is the recession coefficient of the upper response reservoir in the quick runoff routine of the HBV model. The recession coefficient is determined using *ALFA* and two additional parameters *hq* [mm/d] and *khq* [d^{-1}] representing respectively a high flow rate and a recession coefficient at a corresponding reservoir volume [mm] (see Appendix D). In this study KF showed significant simple linear relations with *URBAN* (-0.72), *SAAR* (0.75) and *PWET* (0.73). The forward entry method was executed by taking *SAAR* as the initial variable. By adding other variables to the equation the strength was not improved and the simple relation with R^2 of 56.35% was accepted. The statistical characteristics are shown in Table 5-9.

Table 5-9: *Statistical characteristics for the regression equation KF*

$KF = \beta_0 + \beta_1 \cdot SAAR$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	-0.06555	0.3095	-1.16	0.05636	56.35%
β_1	0.00009	0.0855	2.27	0.00004	

KS

KS is the recession coefficient for the base flow routine in the HBV model and represents the slope of linear behaviour of slow-flow of the catchment with respect to excess water acquired from the quick runoff routine. In the study of Seibert (1999) *KS* is highly correlated with percentage of forest and lake. KS is sensitive to soil type, land use, size and slope of the catchment according to the study done by Hundecha (2004). In the study of Wale (2008) for this area any significant relation between *KS* and catchment characteristics could not be established. In this study KS has significant simple linear relation with *HI* (0.77), *AVGSLOPE* (0.92) and *VER* (-0.88). The forward entry method was executed by adding AVGSLOPE, which is the highest correlated variable, as initial variable. The strength of the equation could not be improved by addition other variables and the simple linear relation was accepted with R^2 of 85.25%. The statistical characteristics are shown in Table 5-10.

Table 5-10: *Statistical characteristics for the regression equation KS*

$KS = \beta_0 + \beta_1 \cdot AVGSLOPE$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	0.0187	0.2093	1.49	0.0125	85.25%
β_1	0.0018	0.0086	4.81	0.0004	

PERC

In the study of Merz and Blöschl (2004) it was shown that *PERC* is negatively related to the river network density and it was suggested that in catchments with few streams a larger portion of water penetrates deep into the subsurface than is the case for catchments with a large river network density. According to Hundecha (2004) percolation is sensitive to soil type. In the study of Wale (2008) a significant simple relation with percentage of Luvisols was established. In this study *PERC* has significant negative relation with *AREA* (-0.82), *LFP* (-0.74) and *DD* (-0.83). By showing the highest negative simple correlation with *DD*, that agreed to the arguments made by Merz and Blöschl (2004). The forward entry method was executed by adding *DD* as the first variable and after including *SAAR*, R^2 increased up to 89.9%. The statistical characteristics are shown in Table 5-11.

Table 5-11: *Statistical characteristics for the regression equation PERC*

$PERC = \beta_0 + \beta_1 \cdot DD + \beta_2 \cdot SAAR$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	7.4926	0.0088	6.11	1.2266	89.9%
β_1	-0.0128	0.0192	-4.61	0.0028	
β_2	-0.0005	0.0864	-2.52	0.0002	

CFLUX

The parameter *CFLUX*, which also belongs to the soil water routine, describes the maximum capillary flow from the upper response box to the soil moisture zone. In this study *CFLUX* has significant simple negative correlation with *SHAPE* (-0.83) and positive correlation with *PDRY* (0.81). Therefore optimisation of the linear relation with the forward entry method was executed with *SHAPE* as the initial variable. The results of the stepwise forward entry regression showed that *CFLUX* is correlated with *SHAPE*, *PDRY* and *PET* with R^2 of 99.8%. The statistical characteristics are shown in Table 5-12.

Table 5-12: *Statistical characteristics for the regression equation CFLUX*

$$CFLUX = \beta_0 + \beta_1 \cdot SHAPE + \beta_2 \cdot PDRY + \beta_3 \cdot PET$$

	Coefficient	*p*-value	t_{cal}	Std error	R^2
β_0	-0.2184	0.2689	-1.52	0.1441	96.27%
β_1	-0.0082	0.0021	-21.86	0.0004	
β_2	0.3867	0.0019	22.63	0.0171	
β_3	0.0007	0.0184	7.28	0.0001	

5.5.3. Validation of the regional model

The regional model is established to predict the discharge from the ungauged catchment. Prior to use, this regional model is assessed by comparing the predicted and observed discharges from the gauged test catchments. It is not possible to carry out a formal validation process as too limited numbers of gauged catchments are available in this study. Therefore validation is done with simulated gauged catchment for the period of 2001 to 2003. Wale (2008) used similar procedure to validate the regional model in his study.

The established regional model was used to estimate the model parameters of gauged catchment using their PCCs. Next the discharge was simulated based on the estimated parameter and the model performance with respect to NS and RVE was evaluated. Table 5-13 shows the parameters derived from the regional model and the model performances.

Table 5-13: *Validation of the regional model of gauged catchments from 2001 to 2003*

	FC	BETA	LP	ALFA	KF	KS	PERC	CFLUX	NS [-]	RVE [%]
Ribb	298.60	1.167	0.699	0.288	0.055	0.098	1.100	0.615	0.85	-1.32
Gilgel Abbay	332.93	1.989	0.720	0.247	0.086	0.084	1.127	1.099	0.83	0.13
Gumara	306.65	1.482	0.827	0.281	0.057	0.079	1.399	0.712	0.75	-22.77
Megech	201.20	1.535	0.700	0.286	0.031	0.085	1.518	0.787	0.54	13.34
Koga	659.31	1.324	0.410	0.425	0.068	0.061	1.626	0.739	0.65	-1.12
Kelti	437.07	2.454	0.724	0.393	0.072	0.054	1.175	1.062	0.53	-42.03

5.5.4. Spatial proximity

The regression method is the most widely used regionalization technique but alternative methods are in use. The choice of catchments from which information is to be transferred is usually based on some sort of similarity measure, i.e. one tends to choose those catchments that are most similar to the site of interest. One common similarity measure is spatial proximity, based on the rationale that catchments that are close to each other will have a similar runoff regime as climate and catchment conditions will only vary smoothly in space (Merz and Blöschl, 2004). Vandewiele and Elias (1995) used a similar approach to estimate parameters of monthly water balance model for 75 catchments from neighboring catchments. In this approach the complete set of model parameters is usually transferred from one or more gauged catchments to ungauged catchments, while in the regression the parameters are regionalized independently from each other. In this study parameter values derived for gauged catchment in upstream areas were transferred to downstream areas of ungauged catchment. Also catchments which are not gauged or failed to be simulated were given parameters from a neighboring catchment. Figure 5-8 shows the catchment relation layout based on the spatial proximity method.

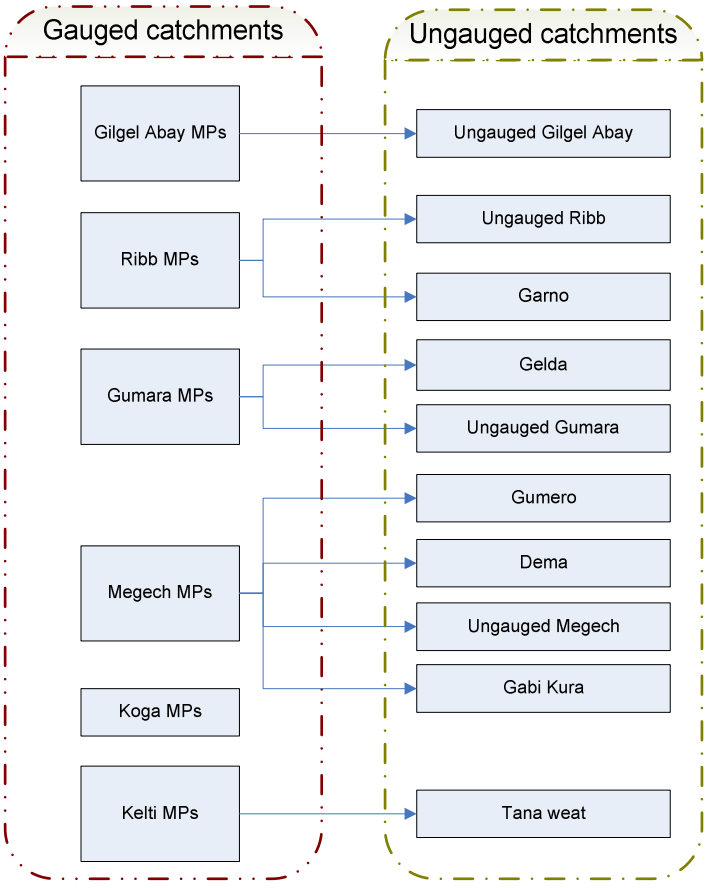


Figure 5-8: *Catchment relation layout in spatial proximity method*

5.5.5. Area ratio

This method considers only catchment areas by assuming that the catchment area is the dominant factor that controls the volume of water as produced by rainfall. The simulated annual average runoff in gauged catchments showed correlation with R^2 72.5% for catchment area (Figure 5-9).

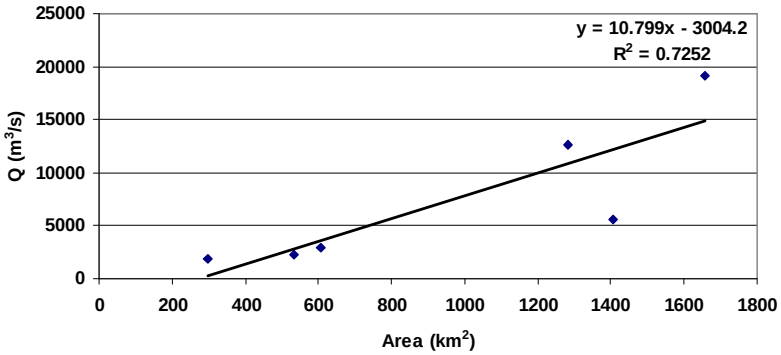


Figure 5-9: *Annual average runoff with respect to catchment area for gauged catchment from 1994 to 2003*

Hence parameter sets of gauged catchments were transferred to ungauged catchments of comparable area. Figure 5-10 shows the catchment relation layout based on the area ratio.

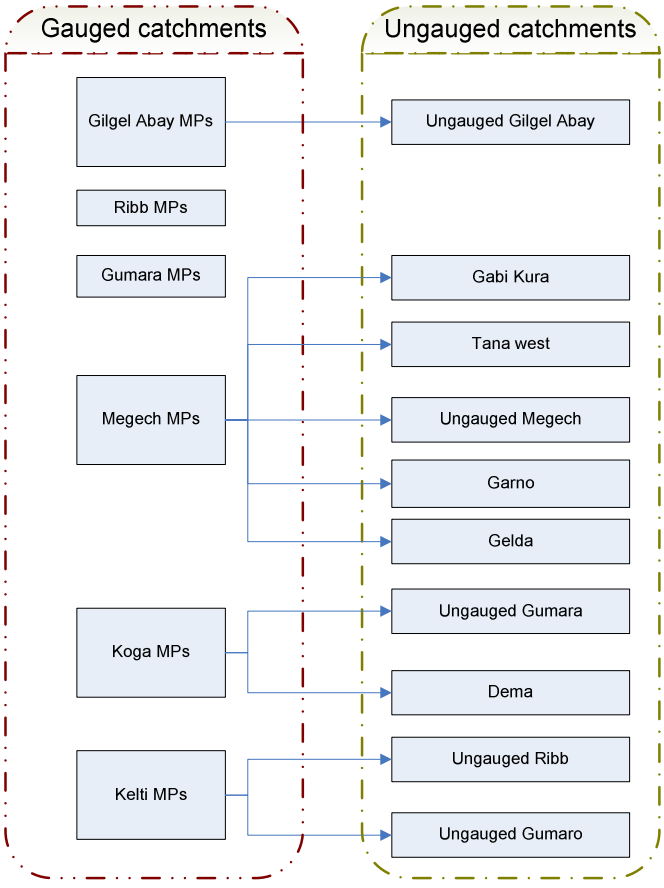


Figure 5-10: *Catchment relation layout based on the area ratio*

5.5.6. Sub-basin mean

The sub-basin mean represents the arithmetic mean (Kim and Kaluarachchi, 2008) of calibrated parameter set of 6 catchments to simulate the flow from ungauged catchment.

6. Simulating lake water balance

6.1. Water balance

A water balance often leads to an understanding of a hydrological system. Several studies were done to study the lake water level simulation and water balance of Lake Tana (Gieske et al., 2008; Kebede et al., 2006; Wale et al., 2008). Dingman (1994) refers to a water balance as "the amount of a conservative quantity entering a control volume during a defined period minus the amount of quantity leaving the control volume during the same period equals the change in the amount of the quantity stored in the control volume during the same time period". The simple water balance equation is formulated as follows:

$$\frac{\Delta S}{\Delta T} = Inflow - Outflow \qquad \textbf{[6-1]}$$

General water balance equation of a lake can be written as:

$$\frac{\Delta S}{\Delta T} = \left(P + Q_{gauged} + Q_{ungauged} + GW_{in}\right) - \left(E + Q_{out} + GW_{out}\right) + ss \qquad \textbf{[6-2]}$$

where, P: is lake areal rainfall, Q_{gauged}: is surface water inflow from gauged catchment, $Q_{ungauged}$: is surface water inflow from ungauged catchment, GW_{in}: is subsurface water inflow, E: is open water evaporation from the lake surface, Q_{out}: is surface water outflow, GW_{out}: is subsurface outflow, ss: is sink source term.

6.2. Water balance terms from observed data

According to the water balance equation there are many processes involved to the lake water balance. Precipitation, inflow from gauged and ungauged catchment and subsurface inflow contribute to the total inflow to the lake. Total outflow from the lake consist of open water evaporation, river outflow from the lake and groundwater recharge. Figure 6-1 depicted the water balance component used in the water level simulation model.

6.2.1. Lake areal rainfall

Seven meteorological stations within and around the lake were selected to calculate the areal rainfall from 1994 to 2003 (see Figure 6-2). Inverse distance weighting functions were used to interpolate the daily observation of these station to be converted to obtain areal coverage. A detailed description of interpolation of rainfall is in Appendix B.

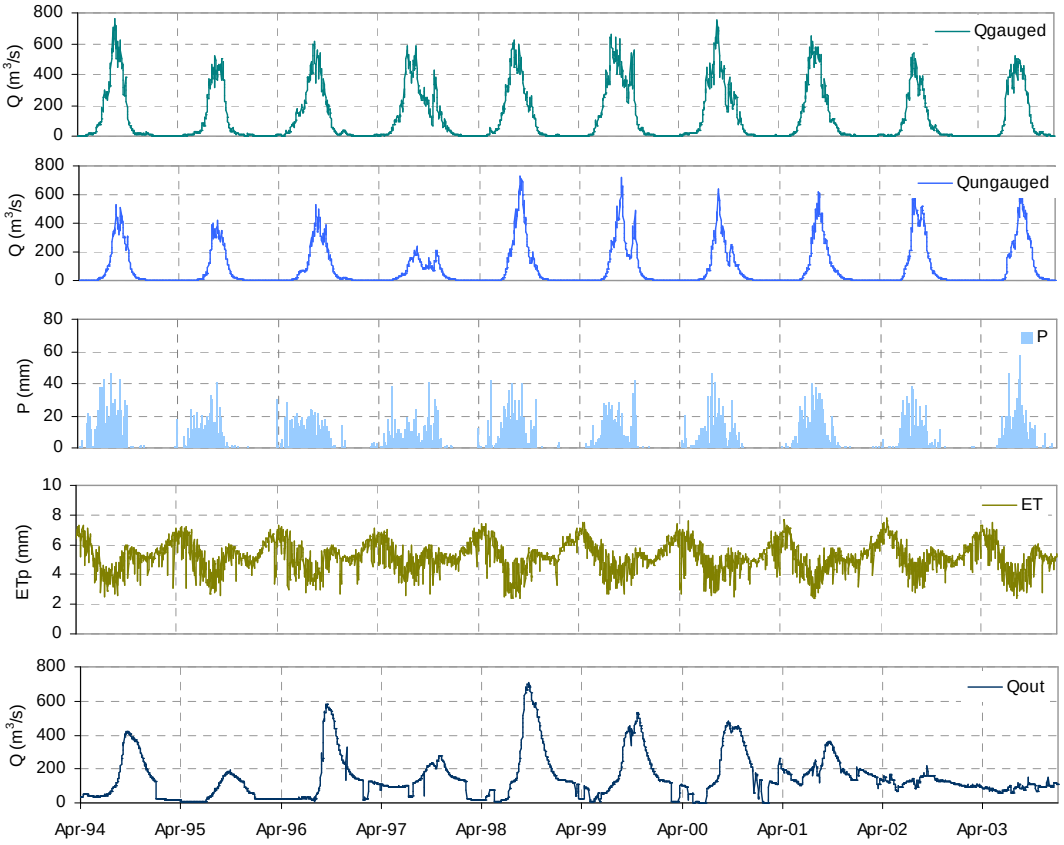


Figure 6-1: *Water balance component used in the water level simulation model*

6.2.2. Open water evaporation

The main factors influencing evaporation from a open water surface are the supply of heat for vaporization and the process of vapour transportation away from the evaporative surface. The solar radiation, wind velocity and the gradient of specific humidity in the air above the open water surface are the major influencing factors. Even though the evaporation is one of the major component of water balance of a lake, it is difficult to estimate. Further direct measurement techniques are not recommended for routine hydrologic engineering applications because they imply a time consuming procedure requiring expensive equipment in order to obtain precise and carefully designed experiments (John, 2006).

In this study the Penman combination approach based on Penman-Monteith method is applied which is widely used as a standard method in hydrological engineering applications to estimate potential evaporation from open water under varying location and climatic condition. The Penman combination approach is formulated as follows:

$$E_p = \frac{\Delta}{\Delta+\gamma} R_n + \frac{\gamma}{\Delta+\gamma} Ea \qquad \textbf{[6-3]}$$

where E_p: is potential evaporation that occurs from the free water evaporation [mm day^{-1}], R_n: is net radiation exchange for the free water surface [mm day^{-1}], Δ: is slope of saturation vapour pressure curve at air temperature [kPa ^{0}C^{-1}] λ: is latent heat of vaporization [MJ kg^{-1}], γ: is psychrometric

constant [kPa $^{0}C^{-1}$], Ea: is the drying power of the air given as a daily rate (mm/day) (Vallet-Coulomb et al., 2001).

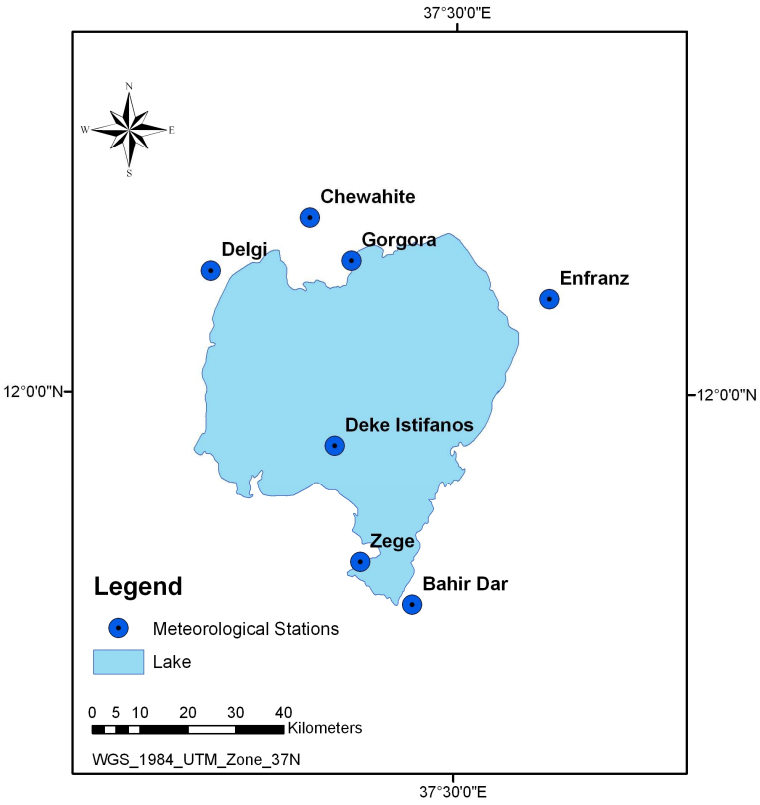


Figure 6-2: *Selected meteorological station in and around Lake Tana to estimate the areal rainfall from 1994 to 2003*

$$E_a = f(u)(e_w - e) \qquad \textbf{[6-4]}$$

where $(e_w - e)$ is the saturation deficit, different between the saturated (e_w) and actual (e) vapour pressure (mbars), f(u) is Penman's function of wind given by:

$$f(u) = 0.26(1 + 0.54U_2) \qquad \textbf{[6-5]}$$

where U_2 is the wind speed measured at 2m height (Vallet-Coulomb et al., 2001).

Penman's equation does not include heat exchange with the ground, and the deep energy loss G (from the water surface into the bottom sediments) and advection term have been ignored here. This assumption is acceptable for monthly or daily estimations in practical hydrological applications (John, 2006). To estimate the net radiation, R_n the following simplified equations are used.

$$R_n = R_s(1-\alpha) - R_L \qquad \textbf{[6-6]}$$

where, R_s: is short wave radiation, α: is surface Albedo, R_L: is long wave radiation

$$R_s = \left(0.25 + 0.5\frac{n}{N}\right)R_a \qquad \textbf{[6-7]}$$

where, N: is maximum possible duration of sunshine hours [hour], n: is actual duration of sunshine [hour], R_a: is extraterrestrial radiation [MJ m^{-2} day^{-1}].

Generally in calculating surface evapotranspiration the value of albedo, which represents the fraction of incoming radiation that is reflected back to the atmosphere, is taken as 0.23 (FAO-56). However river water is much more absorbent and less reflective. This will result in more solar radiation absorbed by the water that as such will warm and evaporate the water. The Albedo can be estimated from a number of satellite sensors and Wale (2008) used landsat ETM+ images to estimate the albedo of the lake within the range of 0.05 to 0.062.

In this study Terra MODIS Level 1 product was acquired for the year 2000 and 2002 from the LAADS Web (http://ladsweb.nascom.nasa.gov/data/search.html) and used Surface Energy Balance System (SEBS) in ILWIS to estimate the albedo. A detailed description of the procedure of estimation of albedo is annexed in Appendix F. Figure 6-3 shows the variation of average albedo during the year lies between 0.08 and 0.16. Other meteorological data was used from the Bahir Dar meteorological station to estimate the open water evaporation by equation [6-3]. The results shows annual average evaporation over the Lake Tana is 1563mm. If the value of albedo is used as 0.06 through out the year the evaporation will be 1670mm/year.

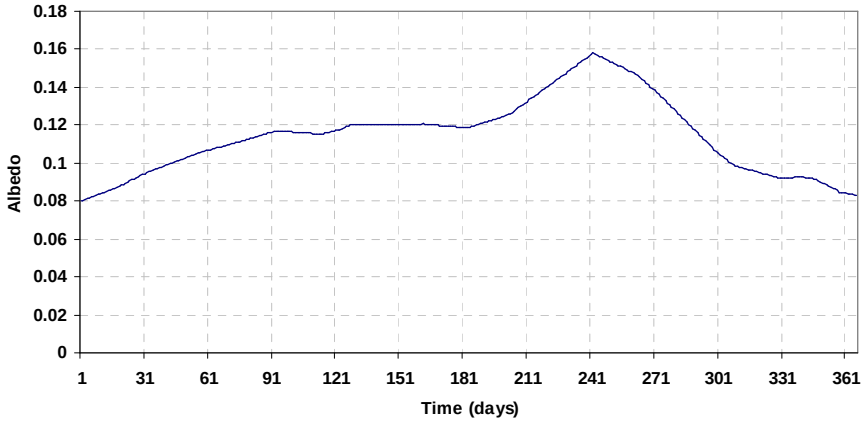


Figure 6-3: *Average albedo variation during the year*

6.2.3. Surface water inflow

The water coming through rivers, streams and over land flows are considered as surface water inflow to the Lake. There are four major gauged rivers dominating surface inflow contributing more than 93% of the total inflow (Kebede et al., 2006; Wale, 2008). The flow data from the gauged catchments with reliable continuous daily flow data are considered directly to the simulation of the water balance and related lake level. The inflow from the ungauged catchment is considered separately for each regionalisation method.

6.2.4. Groundwater inflow and outflow

In the vicinity of the Lake Tana there is no groundwater monitoring data, hence the groundwater flow from the lake and towards the lake is uncertain. However, since the lake is located in a wide depression of plato, it seems to have slight groundwater flow towards the lake. In the report of SMEC (2007) it is mentioned that some 80m thick clay layer underlies the lake floor that prevents inflow or outflow through the lake bottom. By considering the above fact the groundwater component of the water balance is assumed to be negligible and is ignored in this study.

6.2.5. Lake out-flow

Daily water levels of the Abbay station at Bahir Dar were obtained from a database of the Hydrology Department of MoWR and cover the period 1994 to 2003. The daily outflow discharge data is also obtained from the same way for the same period.

6.3. Model development

The above water balance components are integrated according to the equation [6-2] and Lake level is simulated by area-volume and elevation-volume relationship determined by Wale (2008) and Pietangeli (1990) referred by SMEC (2007). A computer code was developed to simulate the water level of the Lake by calculating the water balance component in terms of volume. The initial volume of the lake was fixed according to the observed lake level and the initial lake area was calculated accordingly. This area is used to calculate the lake areal precipitation and open water evaporation as both of them are defined as a function of lake surface area.

$$V_{Lake(i)} = V_{Lake(i-1)} + \Delta S_i \qquad \textbf{[6-8]}$$

$V_{Lake(i)}$ – Lake volume at day i
$V_{Lake(i-1)}$ – Lake volume at day i-1
ΔS_i – Change in storage at day i

$$\frac{\Delta S}{\Delta T} = P + Q_{gauged} + Q_{ungauged} - E - Q_{out} \qquad \textbf{[6-9]}$$

Following equations (6-10 and 6-11) shows the bathymetric relation derived in the previous studies. Polynomial fitted bathymetry by Pietrangeli (1990) reads:

$$E = 1.08\times10^{-9}V^2 + 3.88\times10^{-4}V + 1775.58$$
$$A = 6.2\times10^{-8}V^2 + 1.72\times10^{-2}V + 2516.3 \qquad \textbf{[6-10]}$$

Polynomial fitted bathymetry by Wale (2008) reads:

$$E = 1.21\times10^{-13}V^3 - 1.02\times10^{-8}V^2 + 6.2\times10^{-4}V + 1774.63$$
$$A = 7.93\times10^{-11}V^3 - 5.81\times10^{-6}V^2 + 1.65\times10^{-1}V + 1147.51 \qquad \textbf{[6-11]}$$

The schematic view of the water balance simulation model is shown in Figure 6-4.

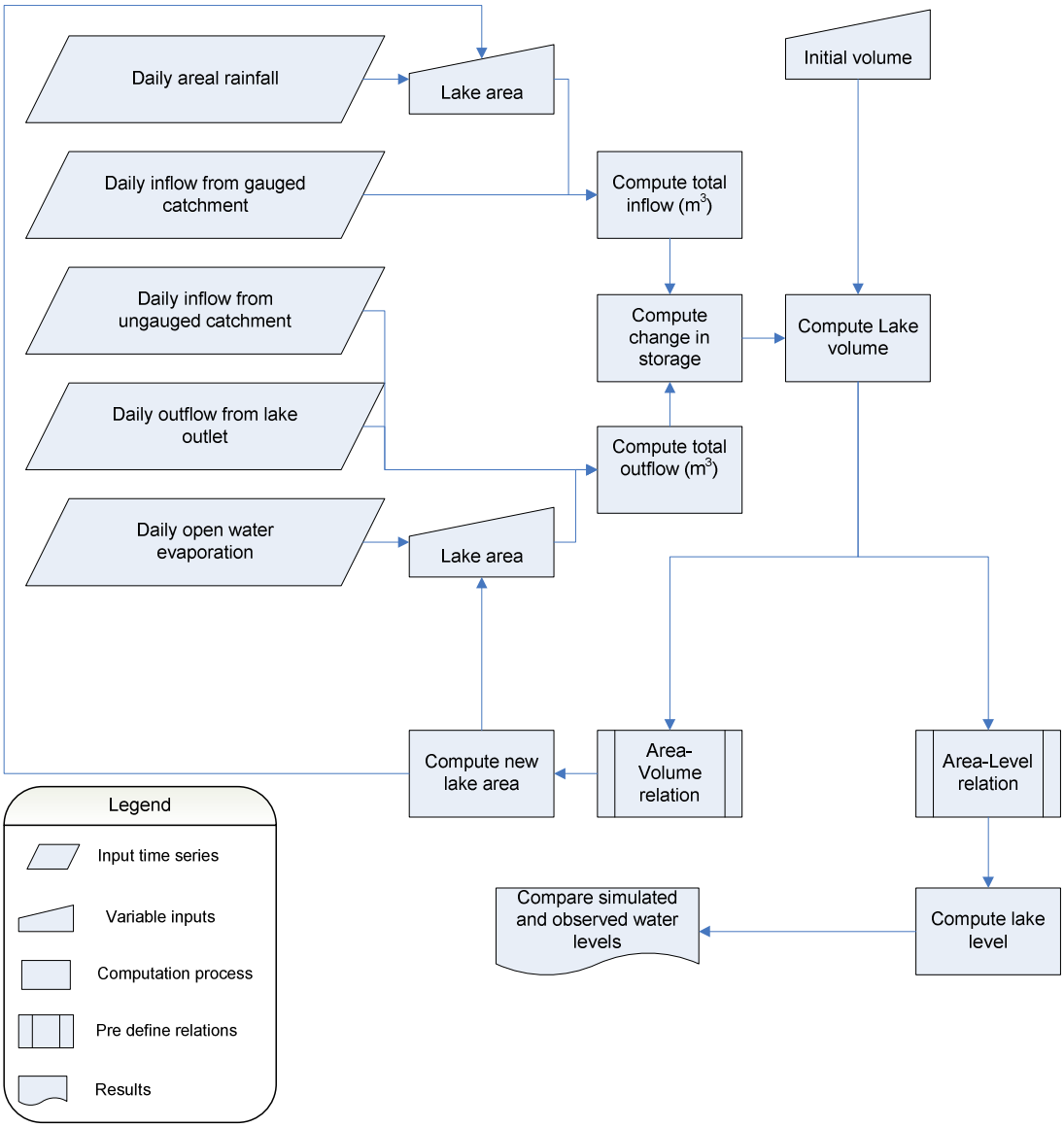


Figure 6-4: *Schematic view of lake water level simulation model*

6.4. Results of water balance

Based on the above mentioned procedure the water level of the Lake Tana was simulated based on the bathymetric relation and the river discharges that estimated from ungauged catchment by regionalization. Figure 6-5 shows comparison of lake level simulation based on Pietrangeli (1990) and Wale (2008). Results are shown in Table 6-1.

Table 6-1: *Results of NS and RVE for selected bathymetric relation*

	NS	RVE (%)
Wale (2008)	0.92	-2.17
Pietrangeli (1990)	0.60	-3.21

These results shows that the bathymetric relation derived by Wale (2008) give good results that as such used to evaluate the reliability of different regionalization techniques.

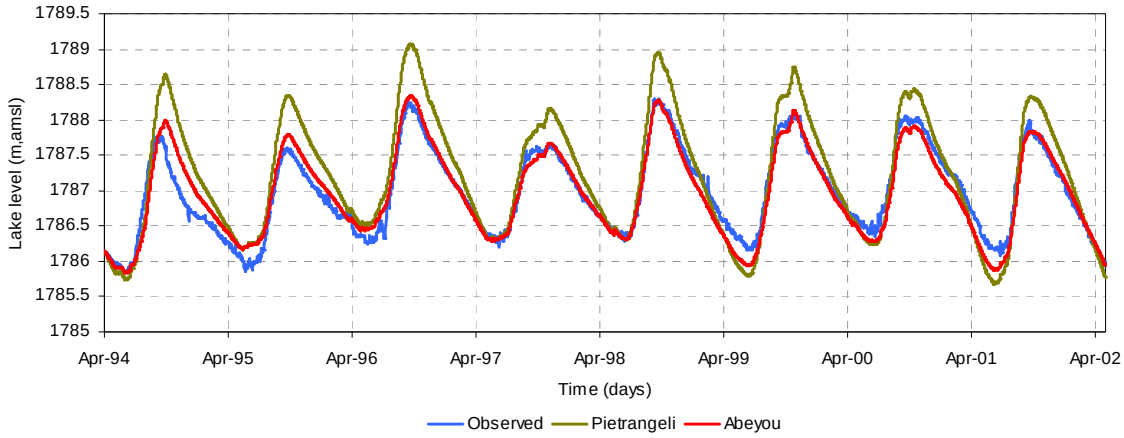


Figure 6-5: *Lake level simulation based on bathymetric relation derived from Pietrangeli (1990) and Wale (2008)*

One of the uncertain components of the Lake Tana water balance is the discharge from ungauged catchments. The discharges of ungauged catchments are estimated using four different regionalization techniques that are used to simulate the water level fluctuation using the result of Wale's interpolated bathymetry (see Figure 6-6). Among the four regionalization techniques, regression method gave the best results and the worst was by sub-basin mean method.

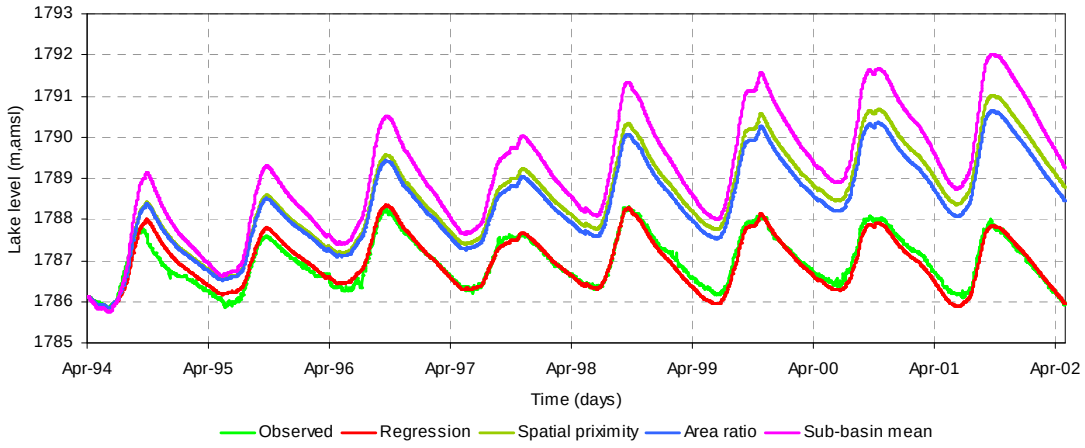


Figure 6-6: *Comparison of simulated lake level in different regionalization techniques*

Except for the regression method, other regionalization techniques show large deviation with observed lake level and simulated, that gradually increase over the simulated period. Hence the result from the regression method is used to calculate the Lake Tana water balance component and results are shown in Table 6-2. It shows that the balance closure term is as large as 85mm. This error of the water balance is 2.7% of the total lake inflow and the lake relative volume error is 2.17%. These errors may be due to uncertainty in lake-groundwater interaction, uncertainty in estimating open water evaporation and lake areal rainfall, runoff from gauged and ungauged catchments.

Table 6-2: *Lake Tana water balance components simulated from 1994 to 2003*

Water balance components	mm/year	MCM/year
Lake areal rainfall	+1347	+4104
Gauged river inflow	+1254	+3821
Ungauged river inflow	+527	+1605
Lake evaporation	-1563	-4762
River outflow	-1480	-4508
Closure term	+85	+260

7. Conclusion and recommendation

7.1. Conclusion

The main objective of this study is predicting discharges at the ungauged catchments as accurate as possible. Therefore, the research has been executed using classical approaches of regionalization techniques with the following steps:

- First best performing model parameter values were identified by calibrating the HBV model against observed discharge in daily time steps for gauged catchments.
- Secondly select regionalization techniques were applied to transfer model parameter to ungauged catchment by
  - Establishing relationship between calibrated model parameters with climatic and physiographic data: jointly call the regional model
  - Transferring model parameters based on the spatial similarity of the catchment
  - Transferring model parameter based on the similarity of area.
  - Taking overall basin average of model parameters.
- Next the model parameter values at the ungauged catchments were estimated.
- Finally daily water level fluctuations of Lake Tana were simulated in order to get insight which regionalization method performs best to predict discharges from ungauged catchments.

To determine the optimum parameter set, an approach was introduced which comprehends application of Monte Carlo simulations using a composed multiple objective function (MOF). In order to assess the model performance, two single objective functions (SOFs) were incorporated in the MOF with each evaluating a particular aspect of the hydrograph. The two selected SOFs are the commonly used Nash-Sutcliffe coefficient (NS) and the Relative Volume Error (RVE).

A review of previous studies was carried out to select appropriate model parameters for calibration since in this study it is tried to establish a robust regional model which addresses all aspects of the hydrograph. As such efficient and effective model parameters have to be selected. Based on the gained information, in total 8 model parameters are selected which are: *FC, BETA, ALFA, LP, KF, KS, PERC* and *CFLUX*. Based on the sensitivity analysis on this study *KS, PERC* and *CFLUX* show not much effect to the model performance with respect to both NS and RVE performance indicators. *FC, BETA* and *LP* show significant effect to the model with respect to both performance indicators. Hence in case of establishing regional model more attention was paid to the sensitive parameters.

Among the 17 river gauging stations in Lake Tana basin, 9 of them have continuous river discharge data for 10 years from 1994 to 2003 and have areal occupancy of 40.2% of the basin. Out of these nine gauged rivers, only six of them have representative river flow data that can be simulated with a reasonable performance of NS value greater than 0.6 and RVE within ±5%. These catchments cover 38.3% of Lake Tana basin.

Four regionalization methods were used to estimate the river discharge from ungauged catchments. Those are regression method, spatial proximity method, catchment area ratio method and sub-basin mean method. Through application of single linear regression analysis and subsequently multiple linear regression analysis, for each model parameter a significant and relatively strong relationship was established. Simulation of lake level shows satisfactory results for discharge predicted by regression method with 2.17% relative volume error and 0.92 NS value. The results of water balance simulation show gauged catchment river inflow of 1254mm/year, ungauged catchment river inflow of 527mm/year, lake areal rainfall of 1347mm/year, lake open water evaporation of 1563mm/year and lake outflow of 1480mm/year. The closure term is obtained as 85mm/year and it is 2.7% of the total annual lake inflow from rainfall and river inflow.

It is observed that effect of land used changes has considerable effect to the peak runoff and it is concluded that intensive forest in a catchment largely affects reduction of flooding event.

7.2. Recommendations

To further enhance the result of the regionalization and the subsequent lake level simulation the following recommendations are formulated.

- It is observed that the parameters PERC and CFLUX in this study do not show significant effect to the model performance and it can be ignored when applying the HBV model to another regionalization study.

- The various causes of rainfall in Ethiopia lead to a wide range of rainfall distributions in space and time. Available rainfall stations however are not well distributed to represent these rainfall events. As such the use of remote sensing data to estimate areal rainfall should be further explored.

- Remote sensing is often claimed to be effective for calibration purposes in hydrological modelling and more specific in stream flow modelling. In this respect to make use of observed soil moisture by remote sensing for model calibration purposes the HBV computer code can easily be adjusted.

- A study of more advanced automatic model calibration techniques is recommended.

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., 1986. An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. Journal of Hydrology, 87(1-2): 45-59.

Abeyou, W., 2008. Hydrological balance of Lake Tana upper Blue Nile basin, Ethiopia, ITC, Enschede, 94 pp.

Arsano, Y. and Tamrat, I., 2005. Ethiopia and the Eastern Nile Basin. Aquatic Sciences 67(1): 15-27.

Bastola, S., Ishidaira, H. and Takeuchi, K., 2008. Regionalisation of hydrological model parameters under parameter uncertainty: A case study involving TOPMODEL and basins across the globe. Journal of Hydrology, 357(3-4): 188-206.

Bergström, S., 1995. The HBV model. In: V.P.Singh (Ed.)Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, CO.

Bergström, S. and Forsman, A., 1973. Development of a conceptual deterministic rainfall - runoff model. In: Nordic hydrology, 4(1973), pp. 147-170.

Bergström, S. and Graham, L.P., 1998. On the scale problem in hydrological modeling. Journal of Hydrology, 211(1-4): 253-265.

Beven, K., 1989. Changing ideas in hydrology -- The case of physically-based models. Journal of Hydrology, 105(1-2): 157-172.

Booij, M.J., 2005. Impact of climate change on river flooding assessed with different spatial model resolutions. Journal of Hydrology, 303: 176-198.

Booij, M.J., Deckers, D.L.E.H., Rientjes, T.H.M. and Krol, M.S., 2007. Regionalization for uncertainty reduction in flows in ungauged basins. In: Quantification and reduction of predictive uncertainty for sustainable water resources management : proceedings of symposium HS2004 at IUGG2007, Perugia, 7-13 July 2007. / ed. by E. Boegh ...[et.al] Wallingford : IAHS, 2007. ISBN 978-1-90150278-09-1(IAHS Publication ; 313) pp. 329-337.

Bras, R.L., 1990. An Introduction to Hydrologic Science. Addison-Wesley Publishing Company, 643 pp.

Conway, D., 2000. The Climate and Hydrology of the Upper Blue Nile River. The Geographical Journal, 166: 49-62.

Conway, D., 2005. From headwater tributaries to international river: Observing and adapting to climate variability and change in the Nile basin. Global Environmental Change Part A, 15(2): 99-114.

Crock, B.F.W. and Norton, J.P., 2000. Regionalisation of Rainfall-Runoff Models. Center for Resorces and Environmental Studies.

Deckers, D.L.E.H., 2006. Predicting discharge at ungauged catchments, University of Twente, Enschede, 143 pp.

Dingman, S.L., 1994. Physical hydrology. Prentice Hall, Englewood Cliffs, NJ.

Dingman, S.L., 2002. Physical hydrology Upper Saddle River. Prentice Hall.

FAO-56, FAO Irrigation and Drainage Paper-No.56.

Fernandez, W., Vogel, R.M. and Sankarasubramanian, A., 2000. Regional calibration of a watershed model. Hydrological Sciences Journal, 45(5): 689-707.

Gheith, H. and Sultan, M., 2002. Construction of a hydrologic model for estimating Wadi runoff and groundwater recharge in the Eastern Desert, Egypt. Journal of Hydrology, 263(1-4): 36-55.

Gieske, A.S.M., Abeyou, W.W., Getachew, H.A., Alemseged, T.H. and Rientjes, T.H.M., 2008. Non-Linear parameterization of lake tana's flow system. Proceedings of the Workshop on "Hydrology and Ecology of the Nile River Basin under Extreame Conditions": 127-144.

Heuvelmans, G., Muys, B. and Feyen, J., 2006. Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets. Journal of Hydrology, 319(1-4): 245-265.

Hundecha, Y. and Bárdossy, A., 2004. Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. Journal of Hydrology, 292(1-4): 281-295.

John, D.V., 2006. Simple version for the Penman evaporation equation using routine weather data. Journal of Hydrology, 331: 690-702.

Kebede, S., Travi, Y., Alemayehu, T. and Marc, V., 2006. Water balance of Lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia. Journal of Hydrology, 316(1-4): 233-247.

Kim, U. and Kaluarachchi, J.J., 2008. Application of Parameter Estimation and Regionalization Methodologies to Ungauged Basins of the Upper Blue Nile River Basin, Ethiopia. Journal of Hydrology (Under review).

Kokkonen, T.S., Jakeman, A.J., Young, P.C. and Koivusalo, H.J., 2003. Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. Hydrological Processes, 17: 2219-2238.

Krysanova, V., Bronstert, A. and Müller-Wohlfeil, D.-I., 1999. Modelling river discharge for large drainage basins: from lumped to distributed approach. Hydrological Sciences Journal, 44: 313-331

Kuczera, G. and Mroczkowski, M., 1998. Assessment of hydrological parameter uncertainty and the worth of multiresponce data. Water Resources Research, 34: 1481-1489.

Lidén, R. and Harlin, J., 2000. Analysis of conceptual rainfall-runoff modelling performance in different climates. Journal of Hydrology, 238(3-4): 231-247.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology, 201(1-4): 272-288.

Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. Journal of Hydrology, 235(3-4): 276-288.

Menzel, L. and Bürger, G., 2002. Climate change scenarios and runoff response in the Mulde catchment (Southern Elbe, Germany). Journal of Hydrology, 267(1-2): 53-64.

Merz, R. and Blöschl, G., 2004. Regionalisation of catchment model parameters. Journal of Hydrology, 287(1-4): 95-123.

Mwakalila, S., 2003. Estimation of stream flows of ungauged catchments for river basin management. Physics and Chemistry of the Earth, Parts A/B/C, 28(20-27): 935-942.

Parajka, J., Merz, r. and Blöschl, G., 2005. A comparision of regionalization methods for catchment model parameters. Hydrological Earth System Science, 9: 157-171.

Reggiani, P. and Rientjes, T.H.M., 2005. Flux parameterization in the representative elementary watershed approzch: Application to a natural basin. Water Resources Research, 41(W04013).

Rientjes, T.H.M., 2007. Modelling in hydrology. ITC Enschede, The Netherlands.

Seibert, J., 1999. Regionalization of parameter for conceptual rainfall-runoff model. Agricultural and Forest Meteorology(98-99): 279-293.

SMEC, 2007. Hydrological study of the Tana-Beles sub-basin, part 1.

SMHI, 2006. Integrated Hydrological Modelling System Manual, Version 5.1.

Vallet-Coulomb, C., Legesse, D., Gasse, F., Travi, Y. and Chernet, T., 2001. Lake evaporation estimates in tropical Africa (Lake Ziway, Ethiopia). Journal of Hydrology, 245(1-4): 1-18.

Vandewiele, G.L. and Elias, A., 1995. Monthly water balance of ungauged catchments obtained by geographical regionalization. Journal of Hydrology, 170(1-4): 277-291.

Wagener, T. and Wheater, H.S., 2006. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. Journal of Hydrology, 320(1-2): 132-154.

Wale, A., 2008. Hydrological balance of Lake Tana upper Blue Nile basin, Ethiopia, ITC, Enschede, 94 pp.

Wale, A., Rientjes, T.H.M., Dost, R.J.J. and Gieske, A., 2008. Hydrological balance of Lake Tana, Upper Blue Nile basin, Ethiopia. Proceedings of the Workshop on "Hydrology and Ecology of the Nile River Basin under Extreame Conditions": 159-180.

WRB, I.W.G., 2007. World Reference Base for Soil Resources 2006, 2nd Edition. World Soil Resources Reports No. 103, FAO, Rome.

Xingnan Zhang, G.L., 1997. Development of an automatic calibration scheme for the HBV hydrological model. Hydrological Processes, 11(12): 1671-1682.

Xu, C.Y., 1999. Estimation of Parameter of a Conceptua; Water Balance Model for Ungauged catchments. Water Resources Management, 13: 353-368.

Xu, C.Y., 2003. Testing the transferability of regression equations derived from small sub-catchment to a large area in cetral Sweden. Hydrological Earth System Science, 7(3): 317-324.

Yu, P.S. and Yang, T.S., 2000. Using synthetic flow duration curves for rainfall-runoff model calibration at ungauged site. Hydrological Processes, 14: 117-133.

List of Acronyms

AAE	Annual average evapotranspiration
ADT	Average daily temperature
ALFA	Parameter defining the non-linearity of the quick runoff reservoir in the HBV model
ALFA	Response box parameter
BETA	Parameter in soil moisture routing in the HBV model
DEM	Digital elevation model
FAO	Food and agriculture organization
F-test	Value which indicates the significance level of the regression equation
HBV	Hydrologiska Byråns Vattenbalansavdelning (Hydrological Bureau Water balance section)
IAHS	International Association of Hydrological Sciences
ID	Inverse distance
ITCZ	Inter-Tropical Convergence Zone
KF	Recession coefficient upper zone
KS	Recession coefficient lower zone
LP	Limit for potential evapotranspiration
MCS	Monte Carlo Simulation
MOF	Multi objective function
MoWR	Ministry of water resources
MP	Model parameter
NMA	National Meteorological Agency
NS	Nash-Sutcliffe coefficient
PCC	Physical catchment characteristics
PE	Potential evapotranspiration
PERC	Percolation from upper to lower zone
PUB	Prediction in ungauged basin
RVE	Relative volume error
SMHI	Swedish meteorological and hydrological institute
SOF	Single objective function
SRTM	Shuttle Radar Topography Mission
T-test	Statistical hypothesis test in which the test statistics has a student’s t distribution if the null hypothesis is true

Appendix A: List of previous studies

Table A-1: *A list of previous studies related to parameter estimation in ungauged basins, after Kim (2008)*

Study	Year	Model[a]	PCC[b]	Method
Jarboe and Haan	1974	WYM	6, 10, 11	Regression
Magette et al.	1976	KWM	2, 6, 8	Regression
Lall and Olds	1987	Mass Balance	1, 3, 4, 5, 6	Regression
Hughes	1989	OSE2	7, 9, 10	Regression
Servat and Dezetter	1993	CREC, GR3	3, 10	Regression
Vandewiele and Elias	1995	MWB	---	Proximity, Transfer
Xu	1999	MWB	7, 9, 10	Regression
Fernandez et al.	2000	Abcd	9, 11	Regression, Regional calibration
Yu and Yang	2000	HBV	1	Flow duration curve
Yokoo et al.	2001	TANK	1, 9, 10, 11	Regression
Kokkonen et al.	2003	IHACRES	2, 5, 7	Mean, Regression, Transfer
Mwakalila	2003	DBM	1, 3, 6, 7, 9, 10	Regression
Xu	2003	NOPEX	9, 10	Regression
Hundecha and Bárdossy	2004	HBV	1, 2, 7, 9, 10	Optimization of transfer function
Merz and Blöschl	2004	HBV	3, 5, 7, 9	Mean, Preset, Regression, Transfer, Proximity
Parajka et al.	2005	HBV	5, 6, 7, 9, 10, 11	Mean, Proximity, Regression, Similarity
Heuvelmans et al.	2006	SWAT	6, 7, 9, 10, 11	Regression, Neural networks
Wagener and Wheater	2006	RRMT	3, 7, 9	Various regressions
Bárdossy	2007	HBV	1, 2, 7, 9, 10	Transfer
Boughton and Chiew	2007	AWBM	3, 4, 5, 10, 11	Regression
Kim and Kaluarachchi	2008	Two layer WB model	1, 2, 3, 4, 5, 6 7, 8, 9, 10	Global mean, Sub-basin mean, Multiple regression, Regional calibration, Aggregate calibration, Volume fraction, Lumped calibration
Bastola	2008	TOPMODEL	1, 2, 3, 5, 6, 7, 11	Multiple linear regression Neural network

[a]See the reference for further details

[b]PCCs selected for the regional relationships in the references. 1=Drainage area, 2=Basin shape, 3=Precipitation, 4=Temperature, 5=Basin elevation, 6=Channel density or length, 7=Basin slope, 8=Channel slope, 9=Soil type, 10=Land use, 11=Geology.

Appendix B: Data assimilation

Precipitation, evapotranspiration and temperature are three input variables for the HBV model. These data are gathered from National Meteorological Agency (NMA) in Ethiopia and ITC archives. The observation stations and the observed durations are shown in Table B-1.

Table B-1: *Meteorological stations and the available time series of observed data*

	Rainfall	RH	Sunshine	Tmax	Tmin	Wind Speed [2m]
Adet	1994-2003	1994-2003	1994-2003	1994-2003	1994-2003	1994-2003
Sekela	1994-2003					
Dangila	1994-2003	1994-2003	1995-2003	1994-2003	1994-2003	1994-2003
BahirDar	1994-2003	1994-2003	1994-2003	1994-2003	1994-2003	1994-2003
Zege	1995-2003			1995-2003	1995-2003	
D_Estifanos	1994-2003			1998-2003	1998-2003	
Delgi	1994-2003					
Aykel	1994-2003	2001-2003		1994-2003	1994-2003	2001-2002
Gorgora	1994-2003					
Gondar	1994-2003					
Enfiraz	1994-2002			2004-2006	2004-2006	
AddisZemen	1997-2006			2004-2006	2004-2006	
D_Tabor	1994-2003	1994-2003	1994-1997	1994-2003	1994-2003	2004-2006
N_Mewicha	1994-2003	1994-2003	1995-2003	1994-2003	1994-2003	1995-2002
			1994-1995, 1997,			
Ayehu	1994-2003	2002-2003	2002-2003	2001-2006	2001-2006	1995-2003
Yetmen	1994-2003			1998-2003	1998-2003	
Enjebara	1994-2006					

Among these observation, data series were prepared for the input variable as per the following procedures since those data sets include many missing and incorrect information.

B.1-Precipitation

Among these stations 16 stations were selected according to the availability of observed data for the period of 1994-2003. Each data series was consisting of missing data for one day to several weeks. Following three methods were used to calculate the missing values for each data series.

B-1.1 Station-average method

In this approach, P_0 is computed as the simple average of the values at the nearby gauges.

$$p_0 = \frac{1}{G}\sum_{g=1}^{G} p_g \qquad \textbf{[B-1]}$$

where. p_0: missing gauged reading,
$p_{1\text{-}g}$: surrounding gauged stations

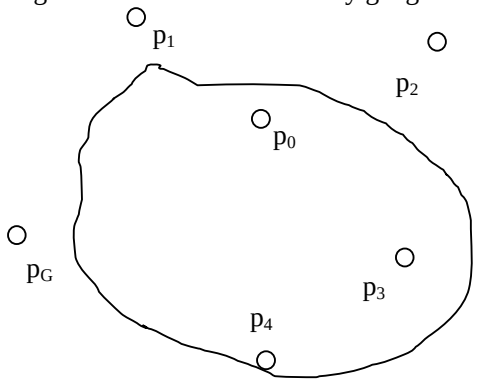


Figure B - 1 *Map showing gauge location*

B-1.2 Normal-ratio method

In this approach the value was estimated at the gauged with missing data by weighting the observations at the G gauges by their respective annual average values, Pg

$$p_0 = \frac{1}{G}\sum_{g=1}^{G} \frac{P_0}{P_g} \cdot p_g \qquad \textbf{[B-2]}$$

where, p0: missing gauged reading, P0: the annual average precipitation at the gauge with missing value, Pg: the annual average precipitation

B-1.3 Regression

Linear regression relation was established with the average rainfall Vs rainfall of the missing station.

$$p_0 = \beta_0 + \beta_1 \overline{p} \qquad \textbf{[B-3]}$$

where. p_0: missing gauged reading, β_0, β_1: regression constant calculated by conventional least-squares regression method, $\overline{p}$: average rainfall of surrounding stations.

B-1.4 Rainfall interpolation

Areal rainfall over the catchments is estimated by inverse distance weighting method. The rainfall intensity at a point P(x,y) out of the rain gauged net work is inversely proportional to distance. The power parameter m controls how the weighting factor reduces as the distance from the reference point increases.

$$\overline{p} = \frac{\sum_{i=1}^{n} \frac{1}{d_i^m} P_s}{\sum_{i=1}^{n} \frac{1}{d_i^m}} \qquad \textbf{[B-4]}$$

where, $\overline{p}$: is areal average rainfall, Ps: is rainfall measured at sub region, di: is distance of station from the region centre, m: is distance weight, n: is number of meteorological stations.

B-2 Evapotranspiration and mean temperature

In order to calculate potential evapotranspiration Penman-Monteith formula was used with variables: maximum and minimum temperature, relative humidity, wind speed and sunshine hours. Among the available meteorological stations 6 stations were selected according to the data availability for the duration of 1994-2003. There were large amount of missing data exist for these four variables. The method described in section A-1.3 was used to establish a regression model for each variable, thus to predict the missing value accordingly. Several combinations were tested to get the highest correlation and that equation was used to calculate the missing values. Figure A-2 showing the relationship between average temperature of surrounding stations and missing station. This experiment was done for the other variables as well and regression model was established in order to calculate the missing values. The outliers were detected by potting the residuals with respect to the mean of the data set. Again these outliers were corrected according to the regression equation.

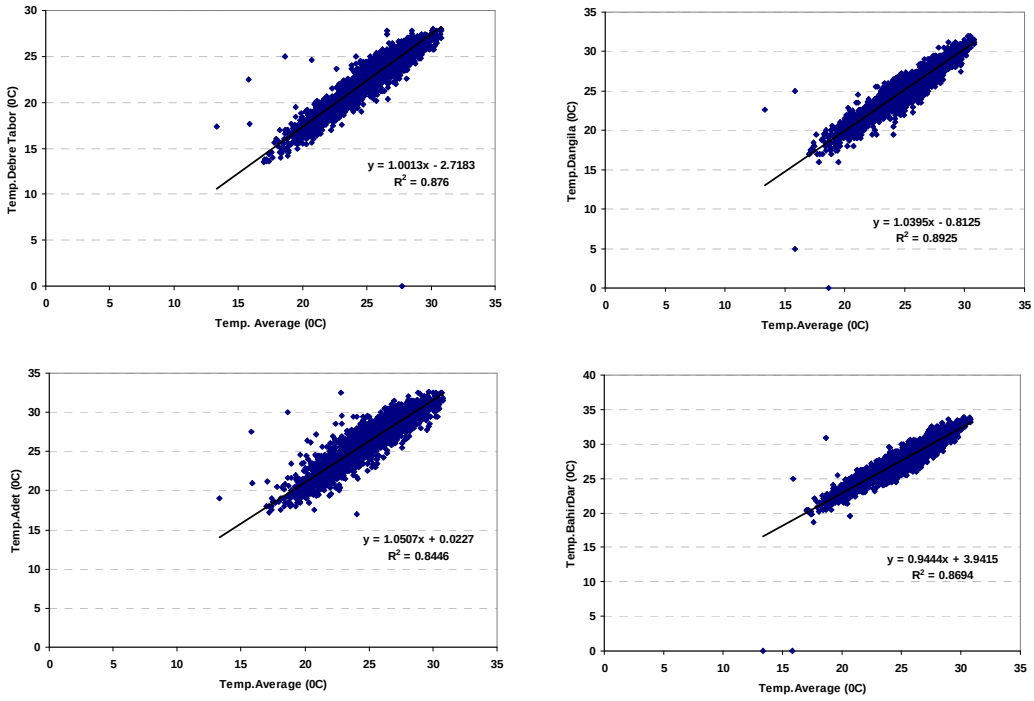


Figure B - 2 *Correlation between average temperature and station temperature*

Mean temperature is calculated by the following equation:

$$T_{mean} = \frac{1}{2}(T_{\max} + T_{\min}) \qquad \text{[B-4]}$$

B-3 Observed discharge data

An important variable for the HBV model is observed discharge data. These discharges are used to calibrate the model parameters. These data sets also consist of missing observations. The following procedures were applied to fill certain gaps of the observed discharge data.

B-3.1 Interpolating discharge gaps for short duration

Unlike rainfall, stream flow shows strong serial correlation; the value on one day is closely related to the value on the previous and following days especially during periods of low flow or recession.

When the gaps in the records are shorts (less than 2 days), during the period of low flows, a linear interpolation technique was used to fill gaps between the last value before the gap and the first value after it. During this operation rainfall values were also examined to check weather there are any significant change occurred during these days. Further the hydrograph of that station was inspected graphically with the neighbouring station to ensure that there are no discontinuities in the flow sequence over the gap.

B-3.2 Interpolating gaps during recession

During the period of recession the flow merely depends on surface and sub-surface storage rather than rainfall. On this period the flow exhibits a pattern of exponential decay curve. When the gaps exist during the long recession periods, interpolation was done between the logarithmically transformed points before and after the gaps.

The slope of the logarithmically transformed flow recession, α (also called a reaction factor) from the last value before the gap Qt0 at time t0 to the first value after the gap Qt1 at time t1 is shown in equation [B-5].

$$\alpha = \frac{\ln Q_{t_o} - \ln Q_{t_1}}{t_1 - t_0} \qquad \text{[B-5]}$$

The reservoir coefficient k is defined as $k=1/\alpha$

Hence discharge Qt at time t within the gap is calculated by equation [B-6]

$$Q_t = Q_{t_o} \exp(-\frac{t - t_0}{k}) \qquad \text{[B-6]}$$

B-3.3 Detecting unreliable and spurious data

Screening was done in order to identify the unreliable and spurious data in the observed discharge time series. The visually identifiable unreliable and spurious data sets were described in section 2.2.4 in Chapter 2. Further analysis is described here to identify the outliers and unreliable individual data points during the high rainfall period.

First the incremental or decremental differences of precipitation (ΔP) and observed discharge (ΔQ) is calculated for each consecutive time steps according to the equation [B-5] and [B-6]. Then the ratio $\frac{|\Delta P|}{\Delta Q}$ was plotted against the time for the entire duration (1994-2003) (Figure B-3 and Figure B-4).

$$\Delta P = P_t - P_{t-1} \qquad \text{[B-5]}$$

$$\Delta Q = Q_t - Q_{t-1} \qquad \text{[B-6]}$$

The majority of value for ratio $\frac{|\Delta P|}{\Delta Q}$ lies within certain interval (this interval may change from basin to basin) commonly the ratios are close to zero, but some points are appeared as outliers. These points are selected and observed in the time series with respect to the rainfall events and adjusted accordingly.

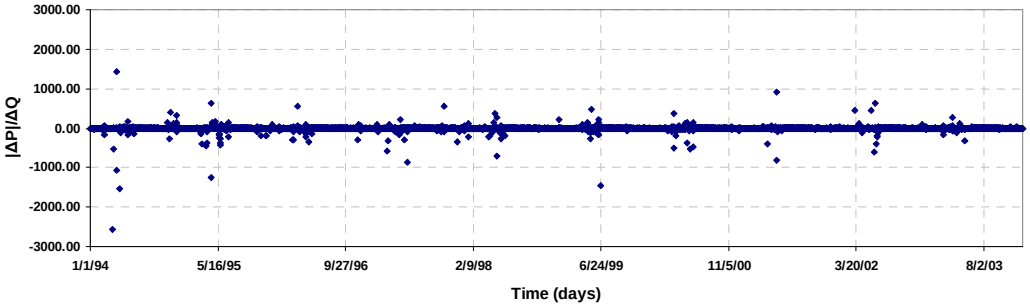


Figure B -3: *Showing the |ΔP|/ΔQ ration during 1994-2003 for Ribb catchment*

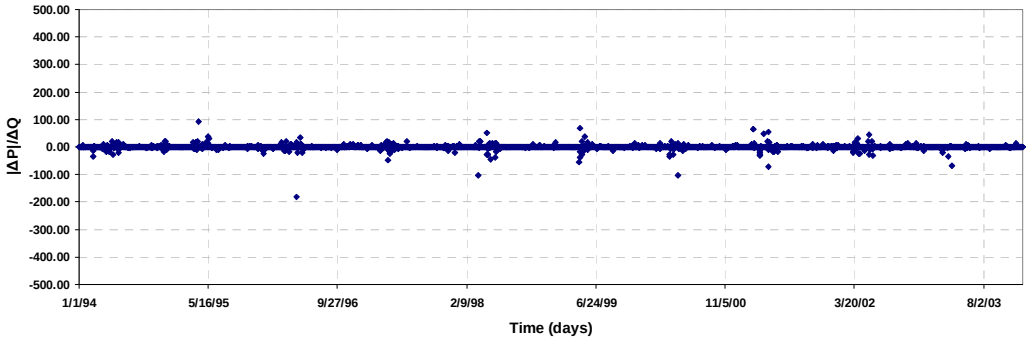


Figure B -4: *Showing the |ΔP|/ΔQ ratio during 1994-2003 for Gilgel Abbay catchment*

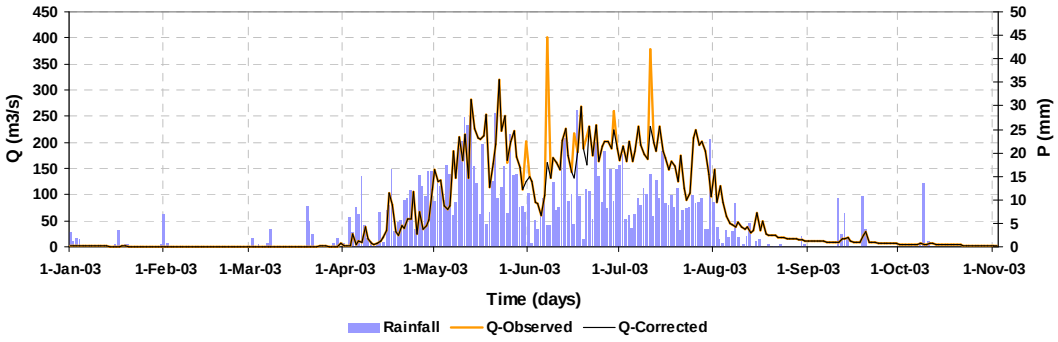
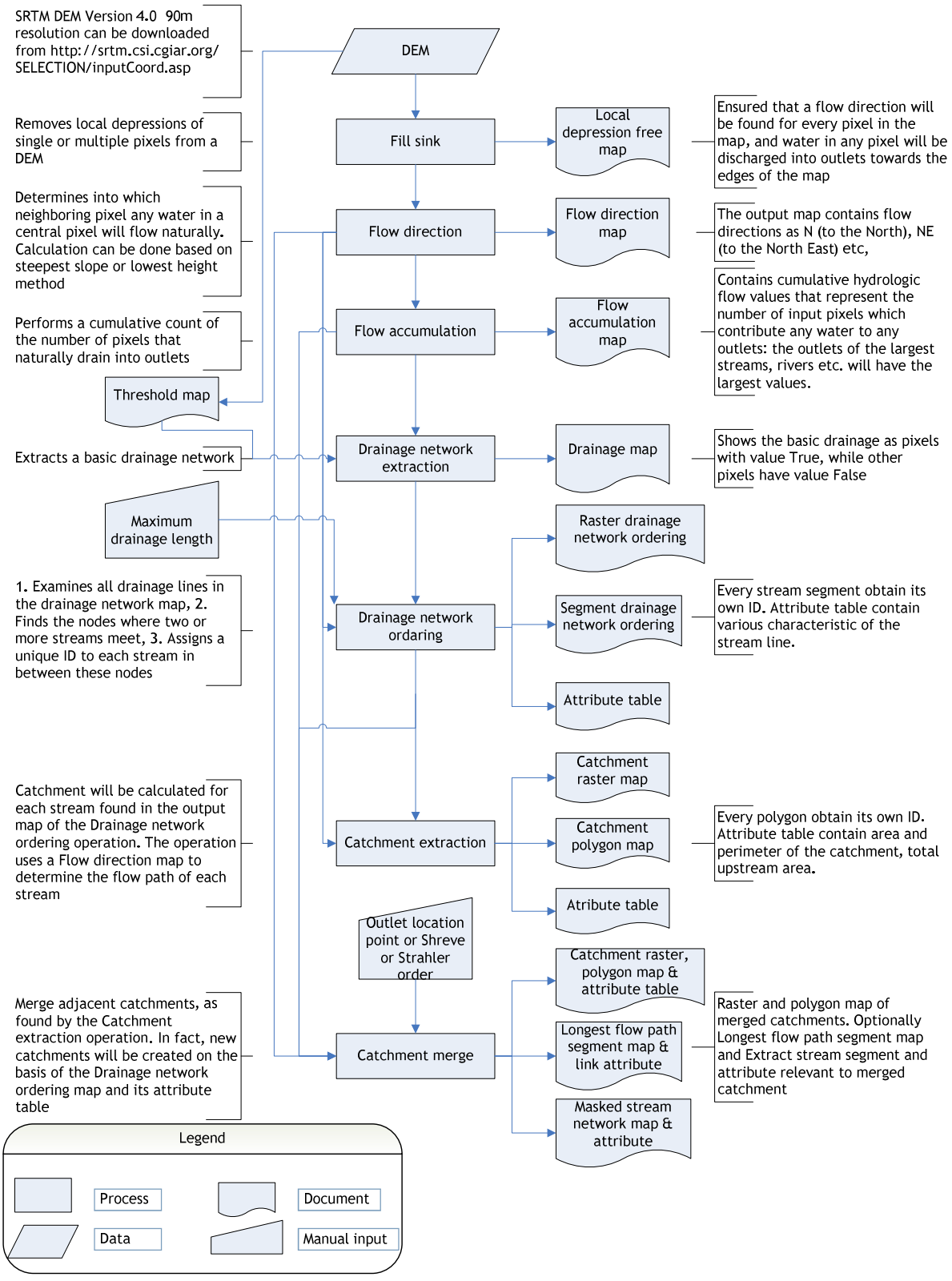


Figure B -5: *Corrected outliers in Gilgel Abbay catchment in the year 2003*

Appendix C: General procedure in DEM hydro processing



Appendix D: HBV model setup

The HBV hydrological model has a long history and the model has found applications in more than 50 countries. Its first application dates back to the early 1970s. Originally the HBV model is developed at the Swedish Meteorological and Hydrological Institute (SMHI) for runoff simulation and hydrological forecasting, but the scope of applications has increased steadily. For instance it has also been used for studies about the effects of climate change. The model has also been subjected to modifications over time, so more specific situations could be addressed.

Experience has shown however, that the standard version of HBV had some major drawbacks which are outlined in Lindström et al. (1997). Therefore a re-evaluation has been carried out and a new model version has been developed. The HBV-96 model is the final result of this model revision (Lindström et al., 1997).

However, the HBV model is not used in the interface released by the SMHI. The computer language FORTRAN and MATHLAB is used and several reasons underpin this decision. When using computer codes, adjustments to the model can be made which are not possible in the regular interface. Adjustments which benefits this study such as the method used for calibrating the model and evaluating the performance of the model. Monte Carlo Simulation (MCS) was embedded in the computer code in order to perform automatic calibration. Previously written computer code by M.J. Booij was adjusted to suit the catchment topology and relevant analysis as the catchment are calibrated individually and no routing routine is imbedded since the catchment are not connected.

The model consists of 6 modules, which are:

1. Precipitation accounting, representing rainfall, snow accumulation and melt;
2. Soil moisture routine, representing actual evapotranspiration;
3. Quick runoff routine, representing quick flow;
4. Base flow routine, representing slow flow;
5. Transformation function, representing slow flow delay and attenuation;
6. Routing routine, representing flow through river reaches.

In HBV model runoff is generated using three meteorological forcing such as precipitation, temperature and potential evapotranspiration. The model's basis is referred to catchment that by themselves can be divided in a number of sub-catchments. The model is semi-distribute, since differences can be made between areas with different altitudes and geographical zones in terms of forest or field.

D-1 Precipitation accounting

Precipitation is one of the main inputs for the HBV model out of the three main variables in simulating runoff processes. Other two variables are mean temperature and potential evapotranspiration. Daily time step of data series were used for all the input variables in this study, but if desirable even smaller time step can also be set.

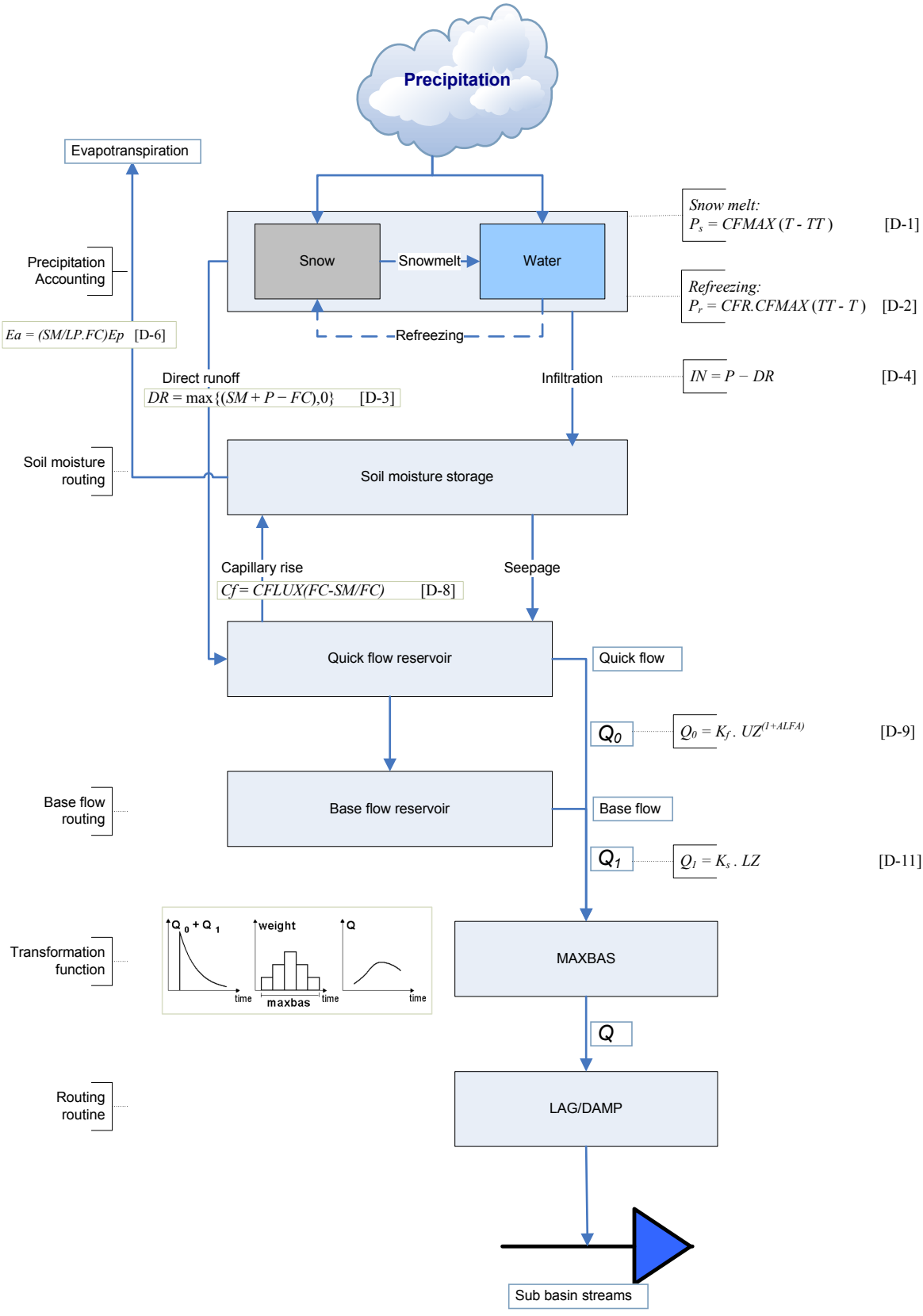

Precipitation
Evapotranspiration
Precipitation Accounting
Ea = (SM/LP.FC)Ep [D-6]
Snow
Snowmelt
Water
Refreezing
Snow melt:
$P_s = CFMAX\,(T - TT)$ [D-1]
Refreezing:
$P_r = CFR.CFMAX\,(TT - T)$ [D-2]
Direct runoff
$DR = \max\{(SM + P - FC),0\}$ [D-3]
Infiltration
$IN = P - DR$ [D-4]
Soil moisture routing
Soil moisture storage
Capillary rise
$Cf = CFLUX(FC-SM/FC)$ [D-8]
Seepage
Quick flow reservoir
Quick flow
Q_0
$Q_0 = K_f \,.\, UZ^{(1+ALFA)}$ [D-9]
Base flow routing
Base flow reservoir
Base flow
Q_1
$Q_1 = K_s \,.\, LZ$ [D-11]
Transformation function
$Q_0 + Q_1$
weight
Q
time
maxbas
MAXBAS
Q
Routing routine
LAG/DAMP
Sub basin streams

Figure D-1: *Schematisation of HBV model structure*

The temperature is used for calculations of snow accumulation and melt, but when desired it can be used to adjust the potential evaporation (Lindström et al, 1997; IHMS, 1997). In order to define precipitation as rainfall or snow, a threshold value is used, TT [^{0}C]. When temperature, T [^{0}C], becomes smaller than the threshold value, rainfall devolves to snow. Interaction between these two components takes place through snowmelt and refreezing, respectively shown in equation [D-1] and [D-2].

$$P_s = CFMAX \times (T - TT) \quad \textbf{[D-1]}$$

$$P_r = CFR \times CFMAX \times (TT - T) \quad \textbf{[D-2]}$$

where, CFMAX: melting factor [mm d^{-1} ^{0}C^{-1}], CFR: refreezing factor [-]

D-2 Soil moisture routine

One of the main parts in the HBV approach is the soil moisture routine that controls the runoff formation. Three main output components are generated from this routine. Those are direct runoff, indirect runoff and actual evapotranspiration. For each sub-catchment the runoff is generated independently as they have specific parameter values for soil moisture accounting procedure and response function.

Direct runoff

The unsaturated top soil layer is represented by the soil moisture reservoir and the volume of the soil moisture, *SM[mm]* in the catchment is computed using the precipitation, *P[mm/d]* as an input which is supplied by the precipitation accounting procedure. Until the maximum soil moisture storage, *FC[mm]*, is not exceeded, the precipitation infiltrates into the soil moisture reservoir. If the maximum soil moisture value is exceeded then precipitation will directly contribute to the runoff, *DR[mm/d]* as shown in equation [D-3]:

$$DR = \max\{(SM + P - FC), 0\} \quad \textbf{[D-3]}$$

The volume of infiltrating water, *IN[mm/d]* is calculated from the output of equation [D-3] is shown by the equation [D-4]

$$IN = P - DR \quad \textbf{[D-4]}$$

Indirect runoff

The infiltrated water generally separated into two components. It can be replenished the soil moisture state or it will seep through the soil layer, which is parameterized by recharge, *R[mm/d]*. This indirect runoff, *R* through the soil layer is determined by the amount of infiltrated water, *IN* and the soil moisture content, *SM* through a power relationship with parameter *BETA[-]*. This relation is shown in equation [D-5].

$$R = IN\left(\frac{SM}{FC}\right)^{BETA} \qquad \textbf{[D-5]}$$

This relation between the parameters state that when there in no infiltration occurs, no indirect discharge is generated. Further while increasing the soil moisture content the indirect discharge is increase and when its come up to field capacity all the infiltration will contribute to the recharge.

Evapotranspiration

In the soil moisture routine actual evapotranspiration, $E_a[mm/d]$ is related to the measured potential evapotranspiration, soil moisture content and parameter *LP* which is a soil moisture limit lies between 0 and 1, beyond that value the actual evapotranspiration occurs at potential rate. This relation is shown in equation [D-6] and [D-7].

$$E_a = \left(\frac{SM}{LP \times FC}\right)E_p \qquad \text{if} \qquad SM < (LP.FC) \qquad \textbf{[D-6]}$$

$$E_a = E_p \qquad \text{if} \qquad SM \geq (LP.FC) \qquad \textbf{[D-7]}$$

D-3 Quick runoff routing

The runoff generation routing is represented as a response function which transform excess water from the soil moisture zone to runoff. In the model structure upper non-linear and lower linear reservoirs are included to represent this response functions. The quick flow and the slow flow are represented by these reservoirs. The upper non-linear reservoir is managed by the quick runoff routine. Generally three components can be distinguished in this reservoir. Those are percolation to the slow flow reservoir, capillary rise back to the soil moisture reservoir and quick runoff.

Percolation

The direct runoff, *DR* and indirect discharge, *R* will entered together to the quick runoff reservoir from which a specific amount transfer to the underlying base flow runoff reservoir is called percolation, *PERC[mm/d]*. Percolation will only occur when there is excess water available in quick runoff reservoir.

Capillary rise

The next component in the quick runoff routine is capillary flow, C_f through which water return again to soil moisture routine due to capillary action in unsaturated zone. This capillary flow depends on the amount of water stored in soil moisture reservoir. The maximum value for capillary flow, *CFLUX[mm/d]* is a parameter in this model and determines the limitation for capillary flow. The capillary flow mainly depends on the soil moisture deficit *(FC-SM)*. There will be no capillary rise if there is no soil moisture deficit. Otherwise, a fraction of the *CFLUX* will flow capillary upward. This is shown in the equation [D-8].

$$C_f = CFLUX.\left(\frac{FC - SM}{FC}\right) \qquad \textbf{[D-8]}$$

Quick runoff

The last component is quick runoff, $Q_0[mm/d]$ which will occur when the recharge from the soil moisture routine is higher than *PERC* and C_f allows and when there is excess water available in the quick runoff reservoir. The Q_0 is determined by the equation [D-9].

$$Q_0 = K_f \cdot UZ^{(1+ALFA)} \quad \textbf{[D-9]}$$

Where *UZ[mm]*: is the storage in the quick runoff reservoir, *ALFA[-]*: is a measurement of the non-linearity of the reservoir, *Kf[1/d]*: is recession coefficient. The recession coefficient is formulated according to the following equation:

$$K_f = \frac{khq^{(1+ALFA)}}{hq^{ALFA}} \quad \textbf{[D-10]}$$

where, parameter *hq[mm/d]* and *khq[1/d]* represent respectively a high flow rate and a recession coefficient at a corresponding reservoir *[mm]*.

D-4 Base flow routine

The base flow routine, which represents the slow flow from a catchment, transforms recharge water from the quick runoff routine. This is represented by equation [D-11].

$$Q_1 = K_s \cdot LZ \quad \textbf{[D-11]}$$

where, $K_s[1/d]$: is the recession coefficient, *LZ[mm]*: is the water level in the reservoir.

D-5 Transformation function

The runoff generated from the response routine ($Q = Q_0 + Q_1$) is routed through a transformation function in order to smooth the shape of the hydrograph at the outlet of the sub-basin. The transformation function is a simple filter technique with a triangular distribution of the weights, as shown in Figure D-2. The generated runoff of one time step is redistributed over the following days using one free parameter (*MAXBAS*). A value of higher than 1 for *MAXBAS* will redistribute the runoff over longer period of time. As a result, this will lead to a delay in the sub-catchment's discharge.

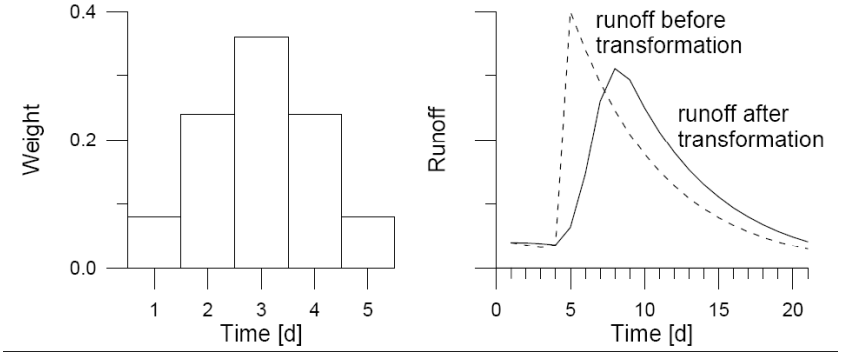


***Figure D-2**: Example of transformation function with MAXBAS = 5 (Seibert, 2002)*

D-6 Routing routine

The routing function also combines runoff discharge for separate sub-catchment. If there is an inflow of water from a sub-catchment this will be added to the local runoff computed by the model. The inflow from other sub-catchments is assumed to flow through a river channel from the outlet of the upstream sub-basin to the outlet of the current sub-basin where the local runoff is added. If there are inflows from several other sub-catchmets, each of them is supposed to flow through its own river channel to the outlet.

Appendix E: Testing significance of regression equation

E-1 Null-hypothesis

In null-hypothesis, f-test is used to assess the significance of the total regression equation.

The general multiple regression equation is shown in equation [E-1].

$$Y' = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + ... \beta_p X_p + \varepsilon \quad \textbf{[E-1]}$$

where Y' : the estimated dependent variable by the regression equation, β_0 : the intercept, $\beta_{1\text{-}p}$: the regression weights, ε : error term.

$H_{0,general}$: $\beta_1 = \beta_2 = \beta_3 = \ldots\ldots = \beta_p = 0$
$H_{a,general}$: at least one regression weight is not equal to zero

$H_{0,general}$ is rejected when

$F > F_{1\text{-}\alpha,p\text{-}1,n\text{-}p}$

where, $F = \frac{MS_{reg}}{MS_{res}}$, $F_{1\text{-}\alpha,p\text{-}1,n\text{-}p}$: can be determine from the F-table.

The mean sums of squared for regression and residual are calculated by dividing the sums of squares through the degree of freedom.

$$MS_{reg} = \frac{SS_{reg}}{p-1} \quad \textbf{[E-2]}$$

$$MS_{res} = \frac{SS_{res}}{n-p} \quad \textbf{[E-3]}$$

The equations for calculating the sum of squares for regression and residuals are shown respectively by [E-4] and [E-5], n: number of observations, p: number of regression coefficient

$$SS_{reg} = \sum_{j}^{n} (Y_j' - \overline{Y})^2 \quad \textbf{[E-4]}$$

$$SS_{res} = \sum_{j}^{n} (Y_j - Y'_j)^2 \quad \textbf{[E-5]}$$

where, SS_{reg}: sum of squares for regression, SS_{res}: sun of squares for residuals, Y_j': estimated value by the regression equation belong to catchment j, Y_j: observed value belongs to catchment j, $\overline{Y}$: average of all observed values. These statistical characteristics are presented in an ANOVA (analysis of variance) table.

In determining $F_{1-\alpha,p-1,n-p}$ for F-table it depends on α which is applied significance level and the degree of freedom, *df* regarding the residual and the regression equation. The number of independent variables attributed in the regression equation will affect this value.

E-2 Specific hypothesis

In specific hypothesis t-test, based on the t-distribution is used to test each of the independent variable, weather they significantly contribute to the total regression equation. A t-distribution occurs when the number of observations is smaller than 30. In this study these conditions are satisfied and the specific hypothesis can be stated as follows:

$H_{0,specific}$: $\beta_1 = 0$
$H_{0,specific}$: $\beta_2 = 0$
$H_{0,specific}$: $\beta_3 = 0$
.
.
$H_{0,specific}$: $\beta_p = 0$

If $|t_c| > \frac{\beta}{SE}$, $H_{0,specific}$ is rejected for the specific regression coefficients:

where, t_c: critical t-value determined from t-table (Davis,2002) depending on the degree of freedom, *df* and chosen significance level, α, SE: standard error which can be taken from the ANOVA table.

E-3 Strength

With respect to the data set, the goodness of fit of the regression equation is measured by r^2, called coefficient of determination is shown in equation [E-6].

$$r^2 = \frac{SS_{reg}}{SS_{tot}} \qquad \textbf{[E-6]}$$

where, SS_{reg}: sum of squared for regression [E-4], SS_{tot}, the total sum of squares is taken from the ANOVA table and calculated from equation [E-7].

$$SS_{tot} = \sum_{j}^{n} (Y_j - \overline{Y})^2 \qquad \textbf{[E-7]}$$

where, SS_{tot}: total sum of squares for regression, Y_j: observed value belongs to catchment j, $\overline{Y}$: average of all observed values.

When r^2 is approaching a value of 1, the regression line is a good estimator of the data. Then the regression equation explains 100% of the total variance present in the data set.

Appendix F: Albedo calculation from MODIS images

Surface Energy Balance System (SEBS) in ILWIS 3.4 Open source provides a set of routine for bio-graphical parameter extraction with the use of satellite earth observation data. These routine derived the evaporation fraction, net radiation and soil heat flux parameters etc, in combination with meteorological information as inputs. This routine was used in this study to derive the albedo of the lake in order to calculate the open water evaporation in water balance study.

F-1 Image acquiring

Terra or Aqua MODIS images can be acquired from Level 1 and Atmosphere Archive and Distribution system (LAADS Web) products search website. (http://ladsweb.nascom.nasa.gov/data/search.html). In this study MODIS Level 1B Calibrated Radiances 250m, 500m and 1km cloud free products and MOD03 Geolocation 1km products for the same day were selected and acquired for the year 2000 and 2002 over the study area. Terra products were selected by concerning the image capturing time and the position with respect to the location of the study area.

F-2 MODIS Level-1B data conversion

Here, ENVI software was used to process the MODIS level-1 products, which were received as HDF file format, into the GeoTIFF format ready for ILWIS importing. First select the desired band to process. By default all the bands in the input file are selected. But in this case we wish to extract band 1 to 7 and Band 31 and 32. After that the Geolocation file was selected and process in order to convert the satellite and solar zenith and azimuth angle bands and height band into GeoTIFF format. Next use the GDAL file-import tool in ILWIS to import the GeoTIFF files to ILWIS. While processing in ENVI, the raw data will converted to reflectance in this case and no need to do pre processing in ILWIS again.

Note: If one uses ModisSwath Tool to process the level-1 product file, the pre-processing routine should be applied to convert the raw to radiance or reflectance. The required calibration coefficient (scale and offset) provided in the HDF header file can be read with the help of 'HDFView' software.

F-3 Atmospheric correction

SMAC tool in SEBS routine in ILWIS is used to correct the atmospheric effect for visible and near visible bands of MODIS images. Figure F-1 shows the required input data for atmospheric effect correction in SMAC. Derivation of atmospheric correction factors are described bellow.

Optical thickness:

The option button can be checked on atmospheric effect correction dialog box if the map of optical thickness (depth) value is available. Otherwise a value can be entered that will be applied to all pixels. The Level 2.0 quality assured Aerosol Optical Thickness (AOT) data can be obtained from the AERONET (AErosol RObotic NETwork) (http://aeronet.gsfc.nasa.gov/new_web/aerosols.html) web site by selecting AOT in AERONET data type. First nearest site to the study area should be selected on the map and then choose the image acquisition date. In this study no site is situated close to the Lake Tana basin, hence a average value was obtained from the surrounding stations (Nairobi, Kibale, DMN_Maine_Soroa and Bodele). Further the records are not available for the particular dates of image acquired on 2000 and the 2002. Therefore average yearly value was computed. In the SMAC

algorithm a value of AOT is used at wave length of 550nm. But in the records AOT is not provided for a wavelength of 550nm. By plotting the given AOT value (see Figure F-2) against wavelength and by polynomial fitting (see Figure 3) required AOT for wavelength of 550nm can be calculated. The unit of the value is in µm, and the value range is between 0.05 and 0.8. The estimate value was 0.31 µm for this study.

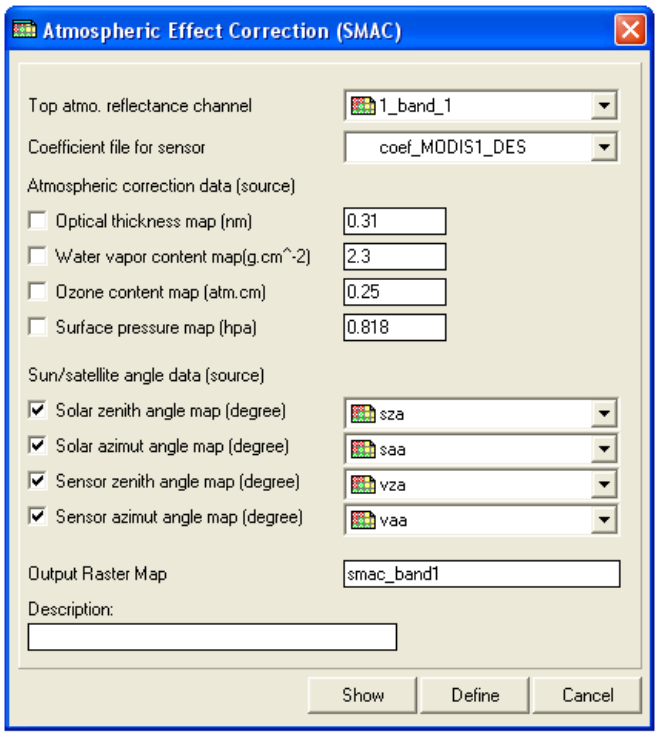


Figure F-1: *Input data for atmospheric effect correction in SMAC tool*

Water vapour content:

The option button can be checked if one wishes to use a map of water pressure content value. Otherwise one value can be given and that will be applied to all the pixels. These values also can be obtained from the above described AERONET web site by selecting 'water vapour' in AERONET data type and choosing the image acquisition date. In this study the average value was computed by selecting surrounding site as there is no nearest site to the study area. The unit of the value is grams cm-2, and the value range is between 0 and 6. The estimated value for this study is 2.3 grams cm-2.

Ozone content:

The option button can be checked if one wishes to use a map containing ozone content value. Otherwise one value can be provided that will be applied to entire image. This value can be obtained by using data from the Ozone Monitoring Instrument (OMI) on the Aura spacecraft at NASA Total Ozone Mapping Spectrometer web site just entering the latitude and longitude and date. (http://jwocky.gsfc.nasa.gov/teacher/ozone_overhead_v8.html).

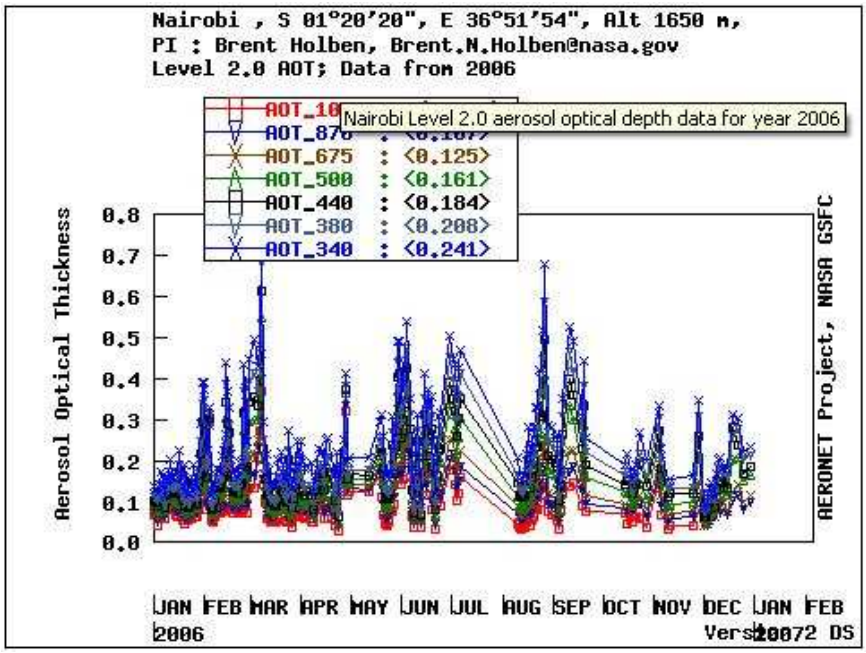


Figure F-2: *Average aerosol optical thickness value for year 2006 in Nairobi*

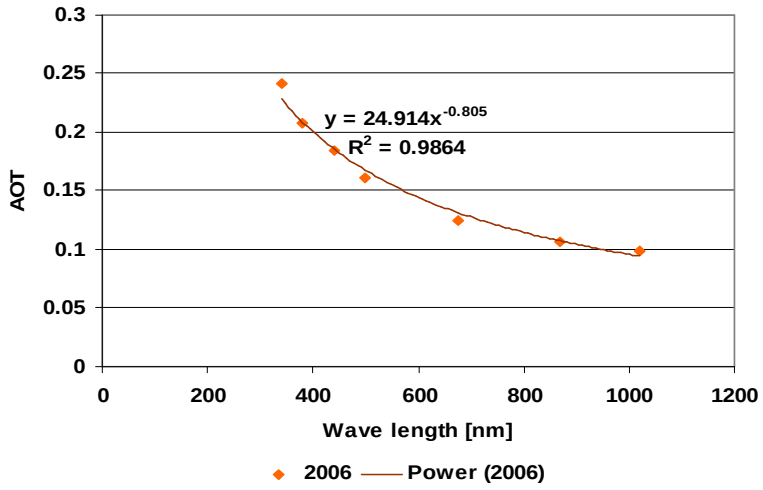


Figure F-3: *The plotting of the AOT measured at different wavelength and polynomial fitting*

Surface pressure:

The option button can be checked if one wishes to use a pressure map. Otherwise one value can be provided that will be applied to entire image. The pressure can be calculated according to the following equation which is employed in FAO 56.

$$P = 101.3\left(\frac{293 - 0.0065z}{293}\right)^{5.26}$$

where P: is the atmospheric pressure [kPa], z: is the elevation above sea level [m]

Solar zenith and azimuth angle map and sensor zenith and azimuth angle map can be acquired with the respective MODIS images. With the use of above factors atmospheric correction can be done for band 1 to 7.

F-4 Land surface albedo computation

The procedure for land surface albedo calculation is as followed:

- Select Operations, SEBS Tools, Pre-processing, Land surface albedo computation. (see Figure F-4)
- Select sensor type MODIS
- Provide the input bands (the atmospheric corrected map in the previous operation)
- Execute the operation by giving the proper output file name.

The albedo over the Lake Tana was calculated for year 200 and 2002 by this procedure (see Figure F-5)

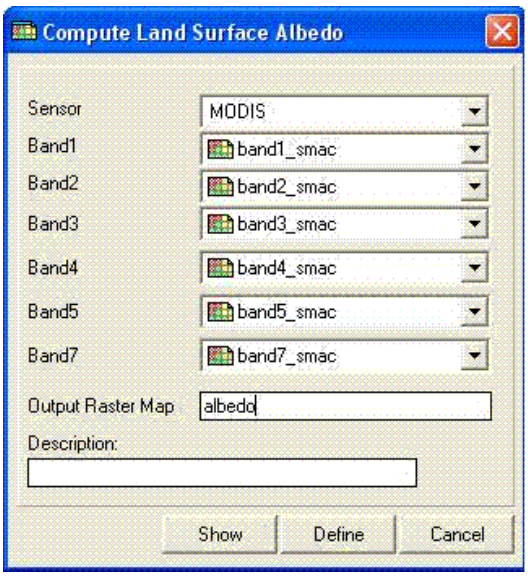


Figure F-4: *Input data for land surface albedo computation*

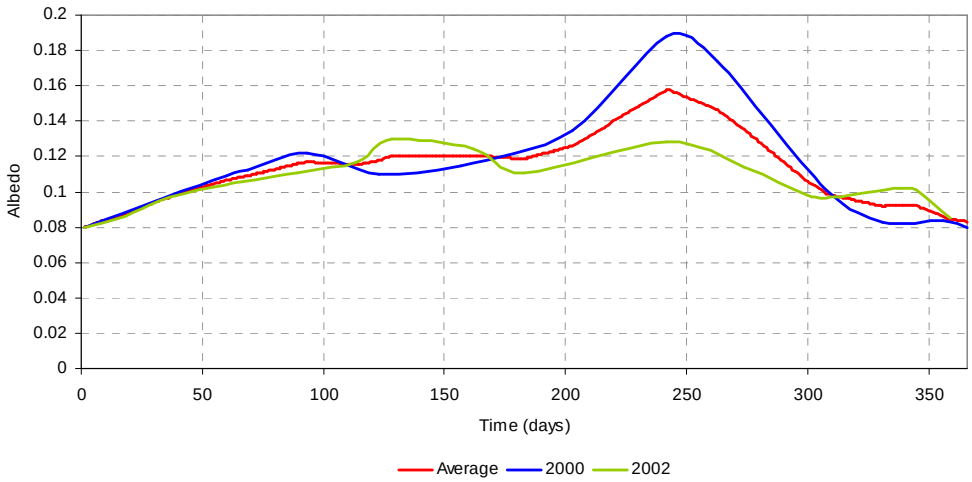


Figure F-5: *Albedo variation over the Lake Tana*

Appendix G: Adjusted parameter space

Name	Description and unit	Prior range	Default value
FC	Maximum soil moisture content [L]	100 - 800	200
BETA	Parameter in soil routine [-]	1 - 4	1
LP	Limit for potential evapotranspiration [LT^{-1}]	0.1 - 1	1
ALFA	Response box parameter [-]	0.1 - 3	0.9
KF	Recession coefficient upper zone [T^{-1}]	0.0005 - 0.15	0.09
KS	Recession coefficient lower zone [T^{-1}]	0.0005 - 0.15	0.01
PERC	Percolation from upper to lower [LT^{-1}]	0.1 - 2.5	0.5
CFLUX	Maximum value of capillary flow [LT^{-1}]	0.0 - 2	0.1

Table G-1: *Selected model parameter and initial parameter space of each parameter*

Parameter	Ribb		Gilgel Abay		Gumara		Megech		Koga		Kelti		Gumero		Garno		Gelda	
	Min	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*
FC	100.0	798.8	135.5	799.9	107.1	799.3	100.0	796.3	101.2	798.9	100.3	799.9	152.6	798.7	104.6	800.0	100.1	695.0
BETA	1.000	3.994	1.005	3.998	1.001	3.997	1.004	3.995	1.001	4.000	1.003	3.995	1.00	4.00	1.01	4.00	1.00	4.00
LP	0.100	0.998	0.103	0.994	0.102	1.000	0.101	1.000	0.10	1.00	0.10	1.00	0.10	1.00	0.10	1.00	0.10	1.00
ALFA	0.101	2.307	0.101	1.964	0.100	2.991	0.106	3.000	0.10	2.99	0.11	3.00	0.10	2.54	0.10	2.97	0.10	1.81
KF	0.0005	0.1499	0.0005	0.1498	0.0005	0.1498	0.0005	0.1500	0.0006	0.1498	0.0008	0.1499	0.0005	0.1499	0.0006	0.1500	0.0006	0.1498
KS	0.0008	0.1500	0.0007	0.1498	0.0009	0.1500	0.0005	0.1496	0.0008	0.1500	0.0014	0.1500	0.0005	0.1499	0.0009	0.1499	0.0012	0.1500
PERC	0.102	2.498	0.101	2.498	0.104	2.500	0.104	2.498	0.106	2.498	0.103	2.498	0.101	2.499	0.106	2.499	0.100	2.500
CFLUX	0.003	1.999	0.000	1.997	0.002	1.999	0.001	1.999	0.003	1.997	0.001	2.000	0.006	1.996	0.002	2.000	0.001	1.993

Table G-2: *Posterior parameter spaces for respective catchment for selected calibration parameter*

Appendix H: Weight of Meteorological station for Gauged and Ungauged catchments

Table H-1: *Weights of rainfall stations by inverse distance squired interpolation for gauged catchments*

	Addis Zemen	Adet	Bahir Dar	Chewahite	Debre Tabor	Dangila	Enfiraz	Gondar	Sekela	Zege	Gorgora	Enjebara
Gelda			0.55		0.16					0.29		
Koga		0.41	0.17			0.15			0.27			
Gumero							0.38	0.44			0.18	
Garno	0.15						0.85					
Kelti						0.94						0.06
Megech				0.06				0.94				
G.Abay						0.27			0.43			0.30
Gumara	0.12				0.88							
Ribb	0.24				0.76							

Table H-2: *Weights of rainfall stations by inverse distance squired interpolation for ungauged catchments*

	Addis Zemen	Bahir Dar	Chewahite	Deke Istifanos	Debre Tabor	Delgi	Enfranz	Gondar	Zege	Gorgora
Ug_Ribb	0.41				0.2		0.39			
Ug_G.Abay		0.27		0.21					0.52	
Ug_Garno							1.0			
Ug_Gelda		0.68							0.32	
Ug_Gumara		0.37		0.29					0.34	
Ug_Gumero							0.26	0.45		0.29
Ug_Megech			0.23					0.58		0.19
Ug_Dema			0.76							0.24
Ug_Tanawest						1.0				
Ug_Gabikura			0.82			0.18				

Table H-3: *Weights of ET stations by inverse distance interpolation for gauged catchments*

	Adet	Bahir Dar	Debre Tabor	Dangila	Gondar	Nefase Mewicha
Gelda		0.65	0.35			
Koga	0.45	0.28		0.27		
Gumero			0.25		0.75	
Garno			0.4		0.6	
Kelti				1.00		
Megech					1.00	
G.Abay	0.40			0.60		
Gumara		0.17	0.65			0.17
Ribb			0.80			0.20

Table H-4: *Weights of ET stations by inverse distance interpolation for ungauged catchments*

	Adet	Bahir Dar	Debre Tabor	Dangila	Gondar
Ungauged Ribb		0.4	0.6		
Ungauged Gilgel.Abbay	0.27	0.43		0.3	
Ungauged Garno			0.47		0.53
Ungauged Gelda		0.78	0.22		
Ungauged Gumara		0.63	0.37		
Ungauged Gumero					1.0
Ungauged Megech					1.0
Ungauged Dema					1.0
Ungauged Tanawest		0.49			0.51
Ungauged Gabikura					1.0

Appendix I: Result of sensitivity analysis

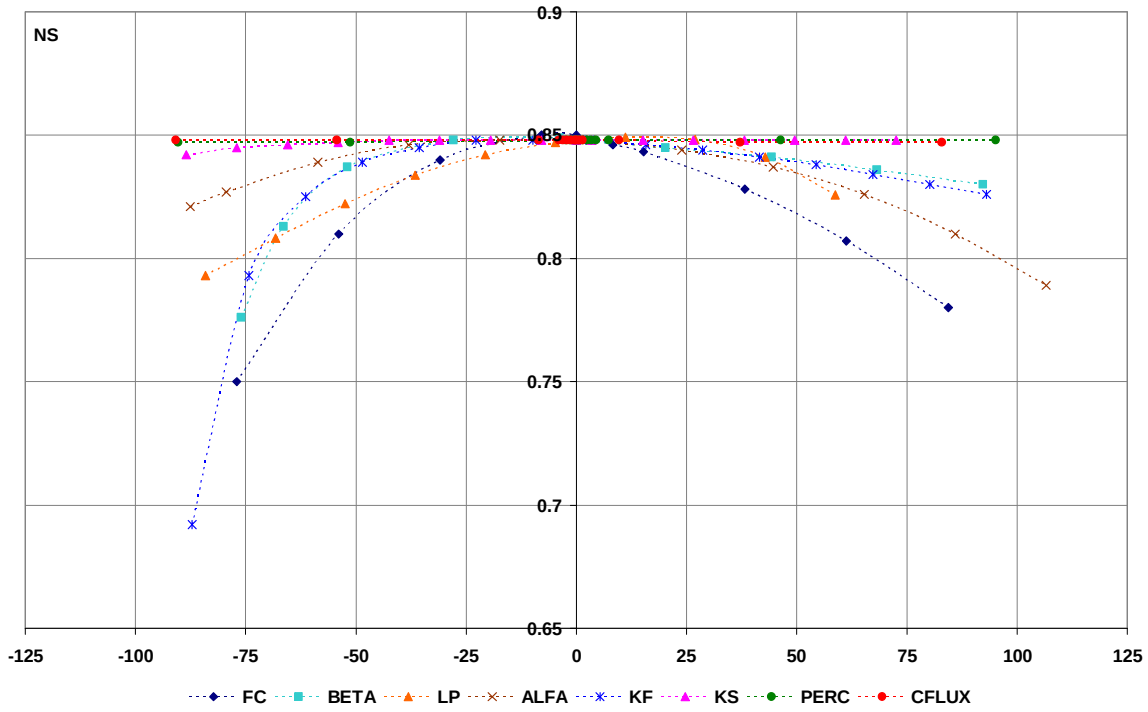


Figure I-1: *Ssensitivity of NS value in Gilgel Abbay catchment*

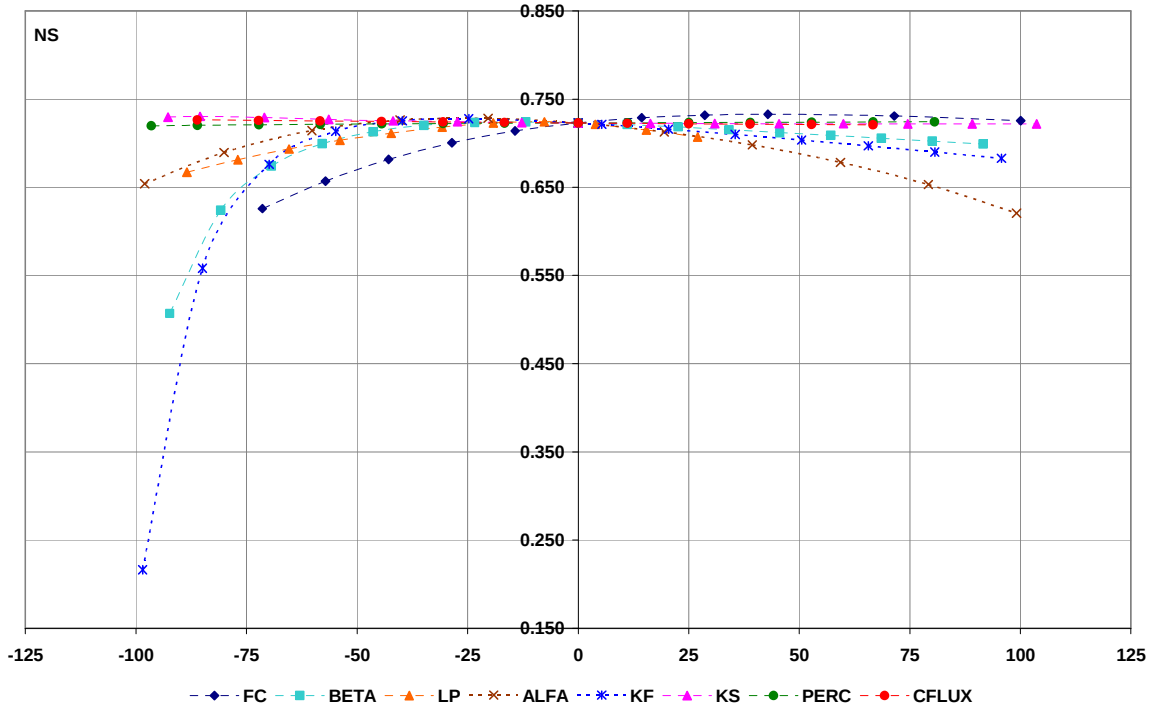


Figure I-2: *Sensitivity of NS value in Gumara catchment*

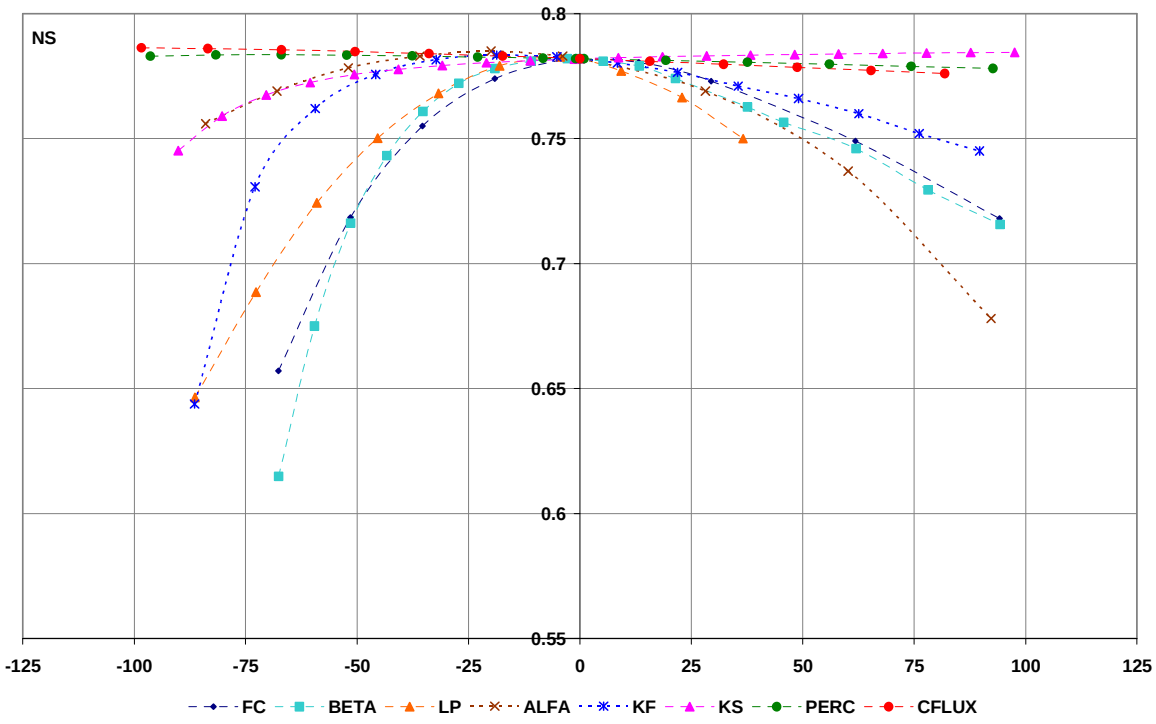


Figure I-3: *Sensitivity of NS value in Ribb catchment*

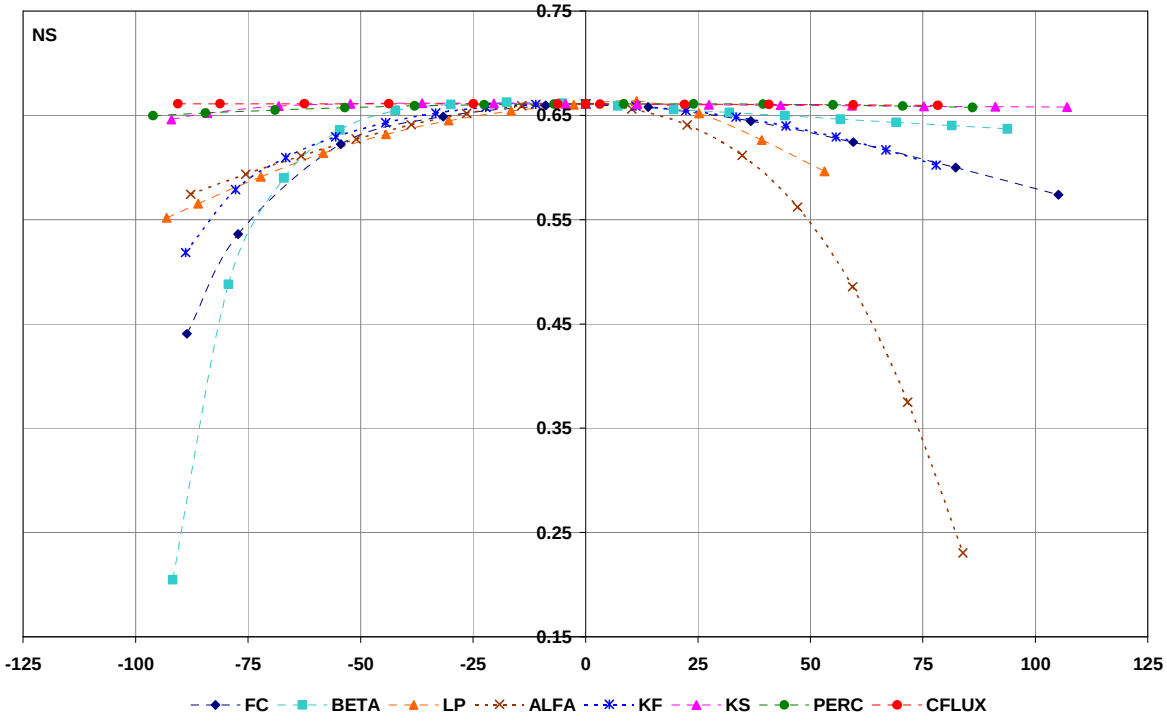


Figure I-4: *Sensitivity of NS value in Kelti catchment*

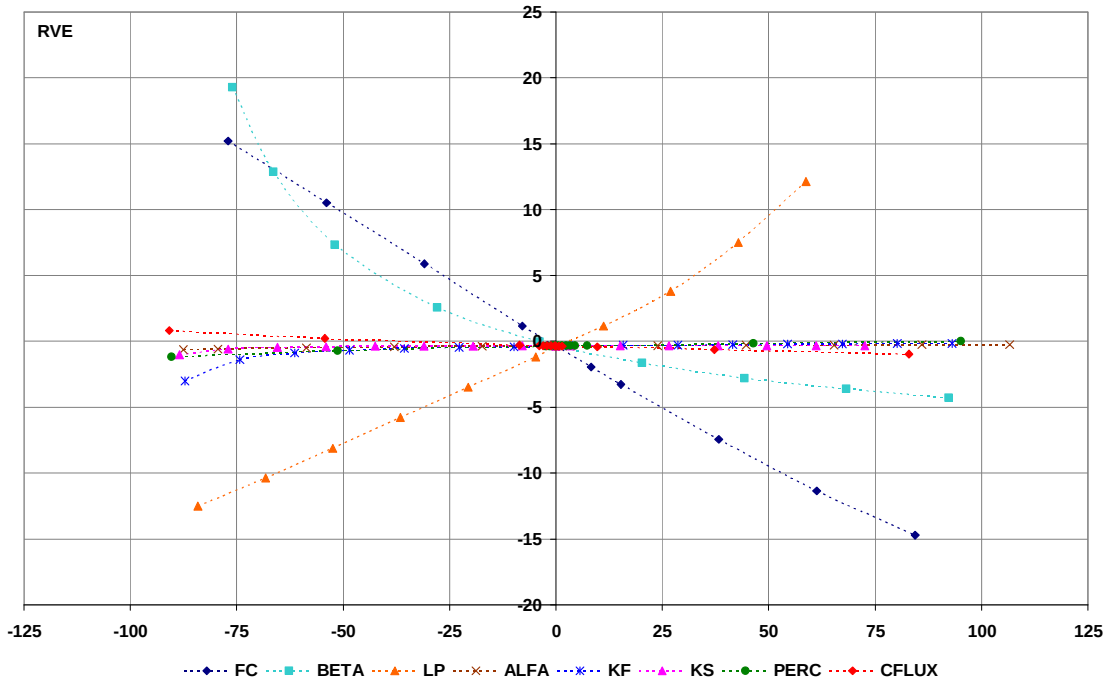


Figure I-5: *Sensitivity of RVE value in Gilgel abbay catchment*

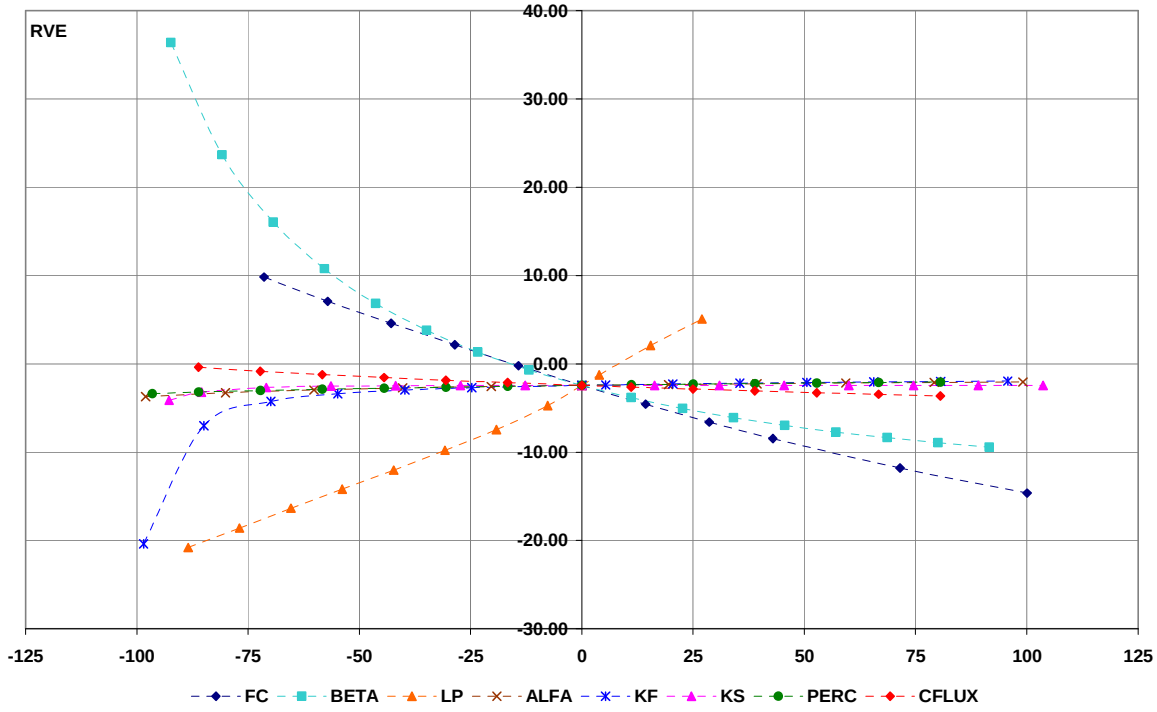


Figure I-6: *Sensitivity of RVE value in Gumara catchment*

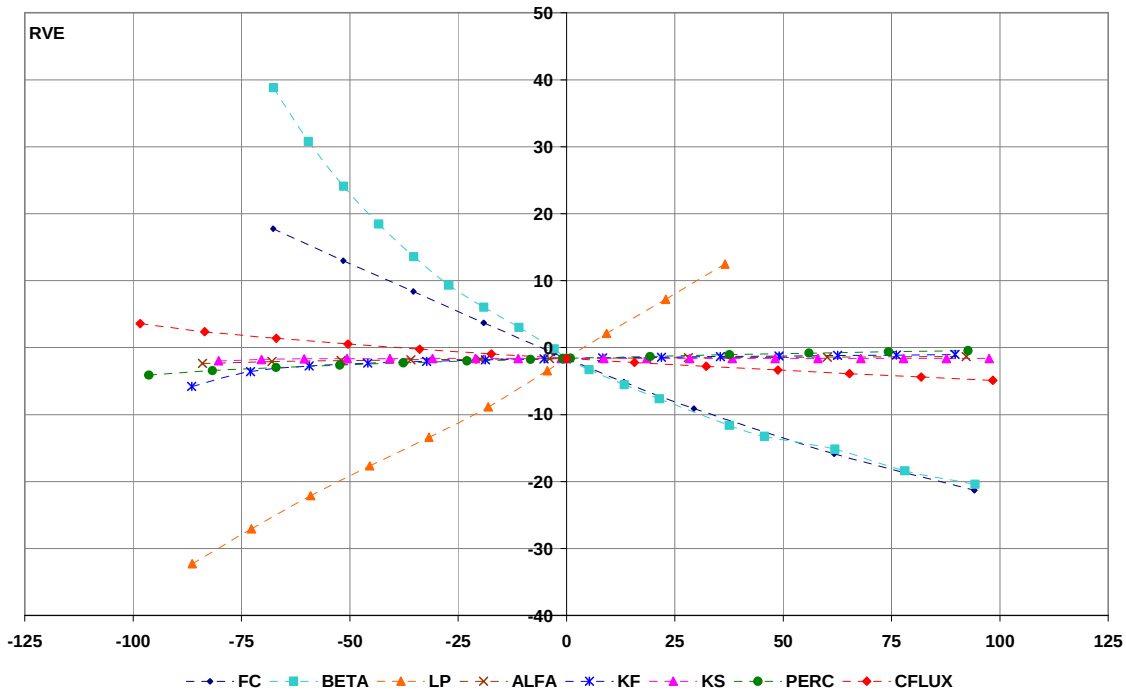


Figure I-6: *Sensitivity of RVE value in Ribb catchment*

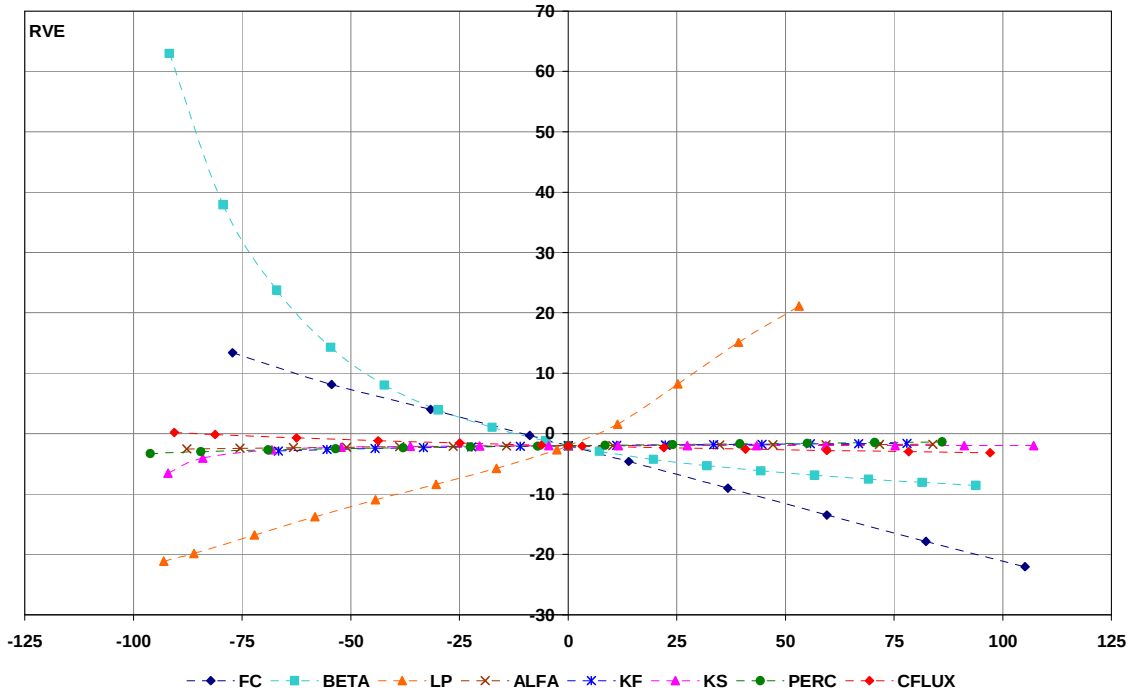


Figure I-6: *Sensitivity of RVE value in Kelti catchment*

Appendix J: The effect of parameter sensitivity

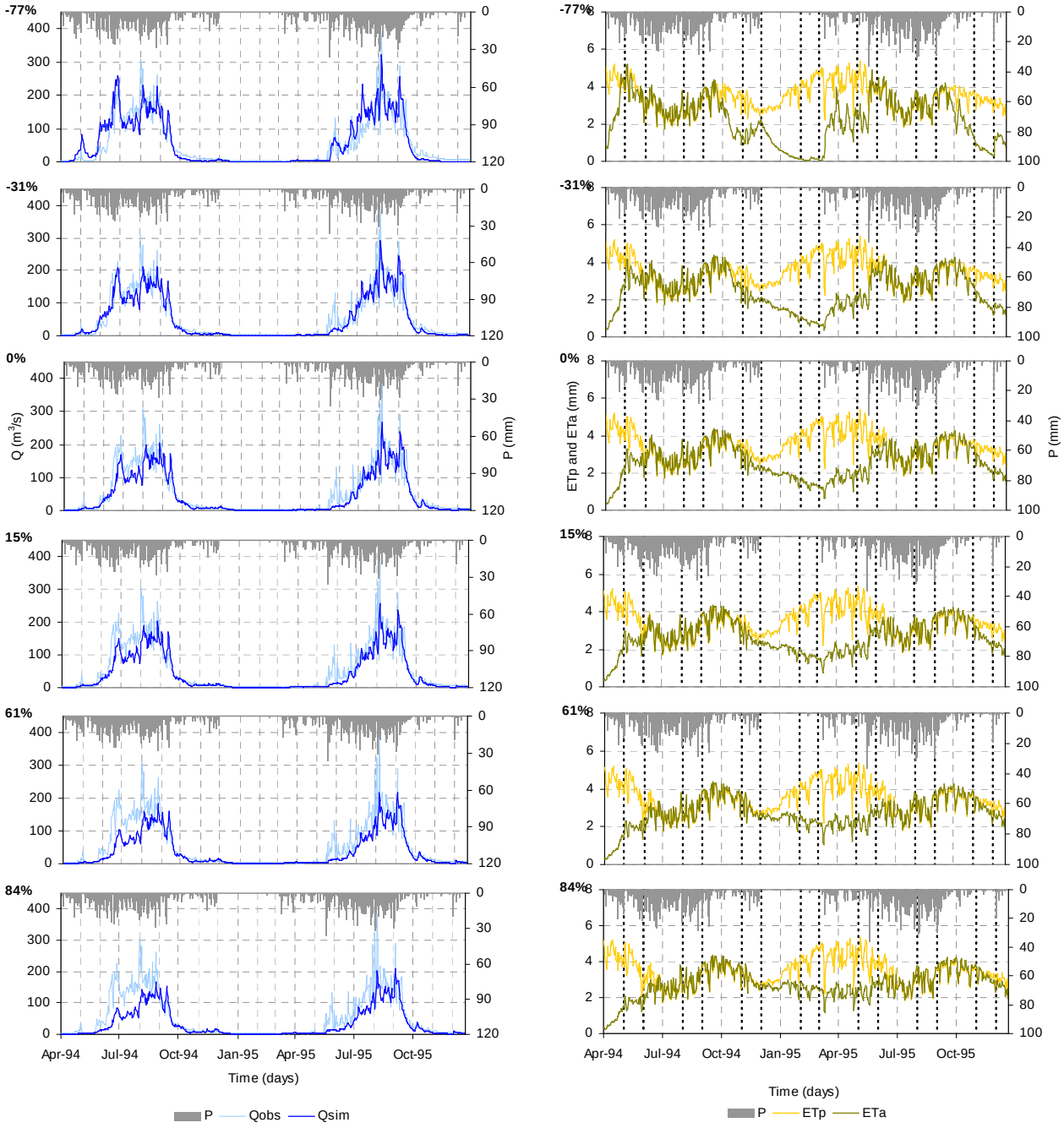


Figure J-1: *Sensitivity of FC for Gilgel Abbay catchment. Left hand side shows variation of simulated hydrograph at each deviation of FC. Right hand side shows the variation of Actual evaporation at each deviation of FC.*

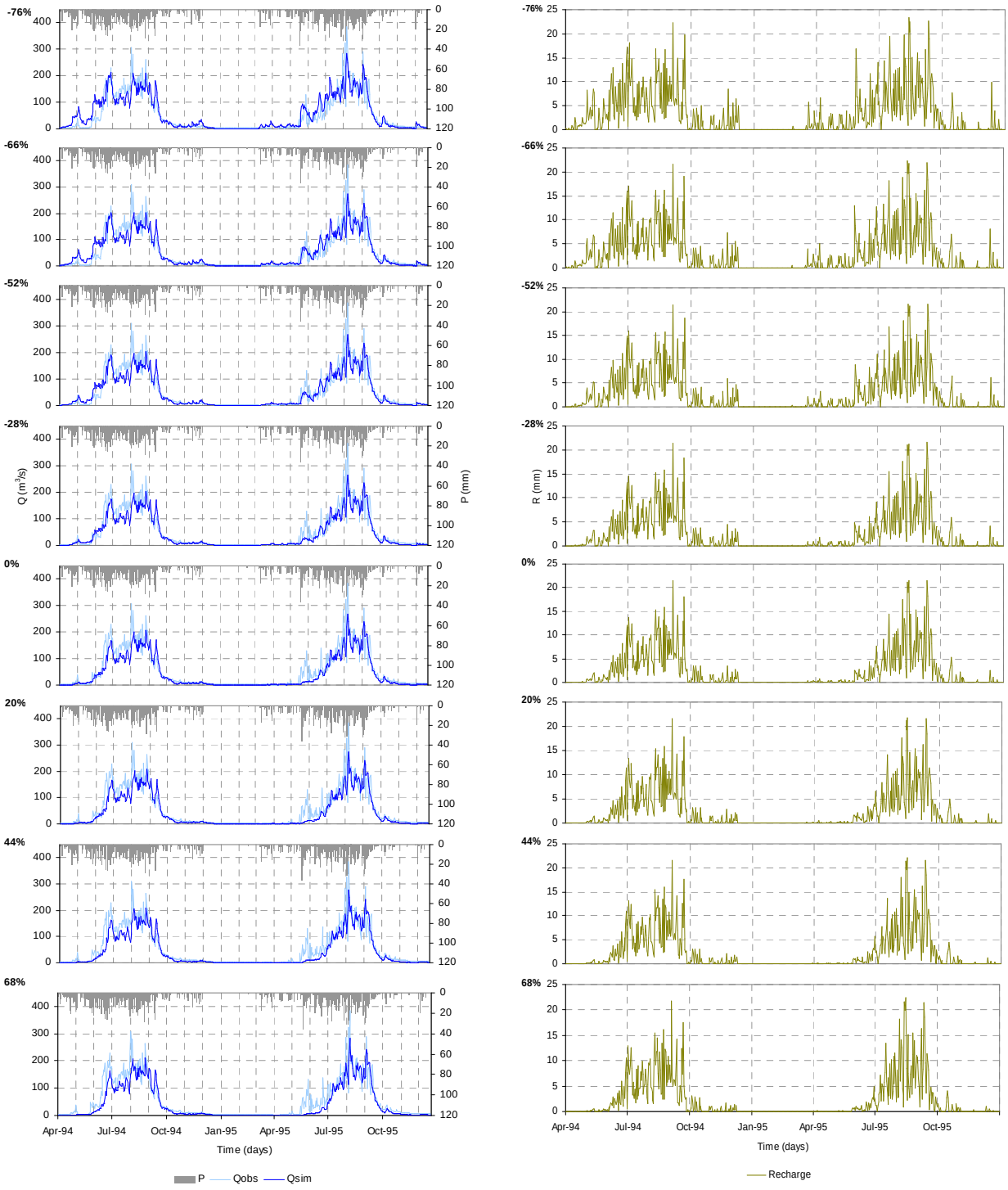


Figure J-2: *Sensitivity of BETA for Gilgel Abbay catchment. Left hand side shows variation of simulated hydrograph at each deviation of BETA. Right hand side shows the variation of recharge at each deviation of BETA.*

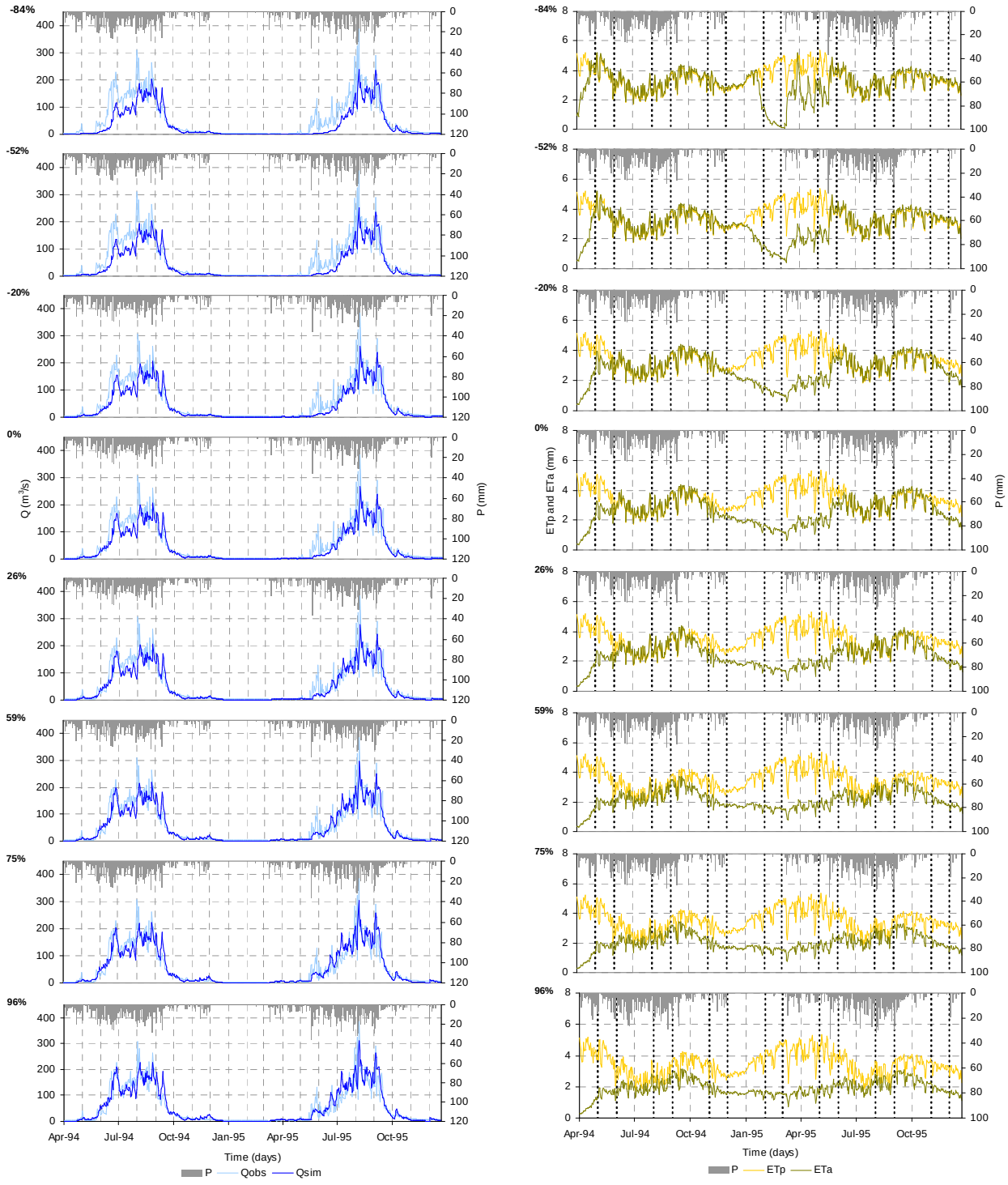


Figure J-3: *Sensitivity of LP for Gilgel Abbay catchment. Left hand side shows variation of simulated hydrograph at each deviation of LP. Right hand side shows the variation of Actual evaporation at each deviation of LP.*

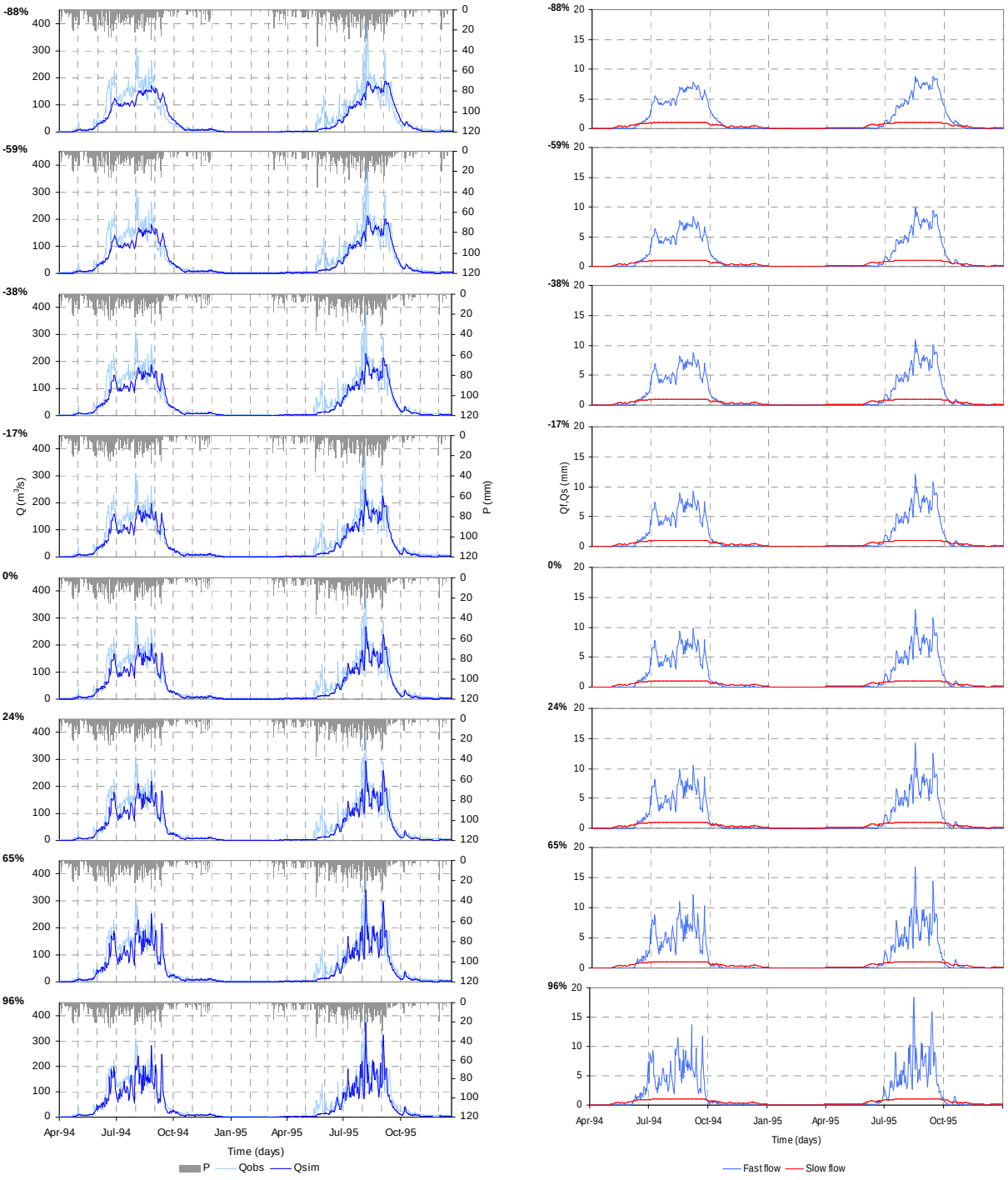


Figure J-4: *Sensitivity of ALFA for Gilgel Abbay catchment. Left hand side shows variation of simulated hydrograph at each deviation of ALFA. Right hand side shows the variation of fast flow and slow flow at each deviation of ALFA.*

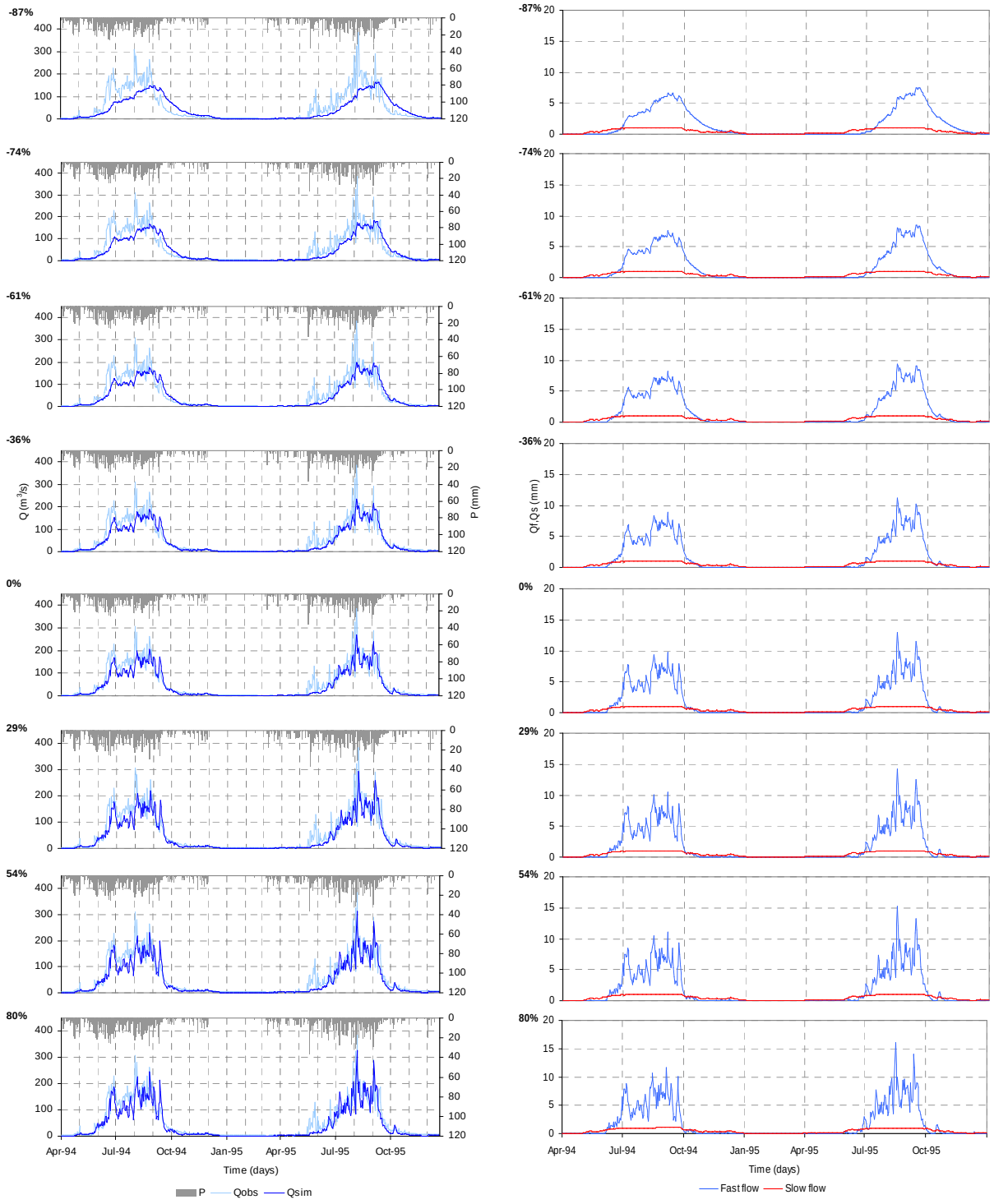


Figure J-4: *Sensitivity of KF for Gilgel Abbay catchment. Left hand side shows variation of simulated hydrograph at each deviation of KF. Right hand side shows the variation of fast flow and slow flow at each deviation of KF.*

Appendix K: Correlation of catchment characteristics

	AREA	LFP	MDEM	HI	AVGSLOPE	SHAPE	DD	CROPD	CROPM	GL	URBAN	FOREST	LEP	NIT	VER	LUV	SAAR	PWET	PDRY	PET
AREA	1.00																			
LFP	**0.90**	1.00																		
MDEM	**0.80**	**0.84**	1.00																	
HI	0.48	0.28	0.35	1.00																
AVGSLOPE	0.67	0.57	**0.83**	0.72	1.00															
SHAPE	-0.16	0.00	0.42	-0.37	0.23	1.00														
DD	0.62	0.68	0.44	0.62	0.57	-0.44	1.00													
CROPD	-0.53	-0.66	**-0.90**	0.03	-0.61	-0.73	-0.10	1.00												
CROPM	0.60	0.61	**0.87**	0.01	0.57	0.66	-0.04	**-0.95**	1.00											
GL	0.48	**0.74**	**0.79**	0.23	0.66	0.41	0.62	-0.75	0.54	1.00										
URBAN	0.22	-0.06	0.21	**0.90**	0.65	-0.15	0.25	0.06	0.03	0.00	1.00									
FOREST	-0.22	0.04	0.28	-0.62	0.03	**0.87**	-0.29	-0.62	0.45	0.46	-0.51	1.00								
LEP	-0.20	-0.28	0.06	0.65	0.54	0.13	0.21	0.06	-0.15	0.19	**0.79**	-0.08	1.00							
NIT	-0.72	**-0.84**	-0.42	-0.10	-0.11	0.40	-0.62	0.23	-0.21	-0.41	0.31	0.24	0.58	1.00						
VER	**-0.79**	-0.68	-0.67	**-0.88**	**-0.85**	0.27	**-0.82**	0.31	-0.29	-0.54	-0.65	0.40	-0.39	0.45	1.00					
LUV	0.18	0.33	-0.13	-0.36	-0.54	-0.37	0.01	0.11	-0.05	-0.11	-0.61	-0.24	**-0.83**	-0.73	0.19	1.00				
SAAR	0.35	0.31	0.00	-0.51	-0.36	-0.25	-0.05	-0.03	0.12	-0.21	-0.66	0.02	**-0.91**	-0.55	0.22	0.61	1.00			
PWET	0.45	0.46	0.16	-0.49	-0.27	-0.16	0.03	-0.19	0.26	-0.04	-0.68	0.10	**-0.92**	-0.66	0.14	0.66	**0.98**	1.00		
PDRY	0.14	-0.01	-0.28	-0.48	-0.49	-0.39	-0.17	0.27	-0.14	-0.50	-0.54	-0.12	**-0.78**	-0.30	0.34	0.46	**0.93**	**0.84**	1.00	
PET	-0.48	-0.73	-0.37	0.35	0.12	0.04	-0.32	0.39	-0.33	-0.45	0.68	-0.20	**0.76**	**0.87**	0.06	**-0.78**	-0.64	**-0.77**	-0.34	1.00

Appendix L: Physical catchment characteristics

	Gauged Catchments	Ribb	G.Abay	Gumara	Megech	Koga	Kelti	Ungauged Catchments	Ribb	G.Abay	Gumara	Megech	Gumero	Garno	Gelda	Dema	Tana West	Gabi Kura
Mean annual evapotranspiration [mm]	PET	1300	1332	1307	1442	1349	1305		1311	1352	1386	1359	1340	1498	1498	1498	1444	1498
Mean precipitation dry season (Oct. to May) [mm]	PDRY	1.10	1.84	1.16	1.00	1.51	1.60		0.84	1.12	0.98	0.96	0.89	0.64	1.02	0.87	0.67	0.79
Mean precipitation wet season (Jun. to Sep.) [mm]	PWET	9.16	10.61	9.21	7.09	9.55	9.74		8.15	9.86	9.13	7.16	7.39	6.96	9.57	7.04	5.33	6.52
Standard annual average rainfall [mm]	SAAR	1395	1750	1415	1117	1542	1585		1210	1486	1368	1162	1120	999	1414	1131	819	992
Luvisol area [%]	LUV	36.84	55.86	87.21	5.03	47.61	91.55		23.94	53.53	54.73	13.69	10.58	1.40	97.52	6.01	34.20	6.99
Vertisol area [%]	VER	0.00	1.95	3.27	3.34	14.67	8.25		13.37	13.26	23.00	47.66	35.17	19.20	1.44	68.57	19.24	64.38
Nitosol area [%]	NIT	0.30	0.51	0.00	9.25	7.69	0.00		0.00	3.13	0.00	0.00	3.65	9.51	0.00	0.00	0.00	0.00
Leptosol area [%]	LEP	38.58	0.00	9.23	81.59	4.11	0.20		9.86	22.15	0.00	18.88	48.65	64.86	0.46	22.64	0.00	1.22
Build up area [%]	FOREST	4.96	0.00	0.45	0.00	6.11	0.00		1.47	4.03	0.00	0.00	0.00	0.54	0.00	0.00	0.25	0.00
Woody savannah [%]	URBAN	0.06	0.08	0.08	0.16	0.00	0.01		0.02	0.12	0.02	0.03	0.01	0.04	0.33	0.02	0.02	0.02
Bare land [%]	GL	12.79	0.00	3.86	0.07	0.03	0.00		8.74	0.04	0.03	0.00	0.00	18.62	0.00	0.00	0.00	0.00
Crop [%]	CROPM	30.66	25.66	31.34	10.23	24.25	0.00		16.06	19.20	19.49	2.06	28.14	7.91	38.93	2.14	33.16	4.13
Grass land [%]	CROPD	51.53	74.26	64.27	89.54	69.61	99.99		73.71	76.60	80.46	97.91	71.85	72.90	60.73	97.84	66.57	95.85
Forest [%]	DD	444.38	428.28	420.27	422.69	397.48	431.05		461.27	389.33	380.00	457.42	395.47	348.10	355.90	397.78	365.77	354.63
Elongation ratio [-]	EL	1.55	1.08	1.50	1.24	1.75	1.39		1.36	1.21	1.25	1.75	1.03	0.85	1.27	2.26	4.69	4.80
Circularity index [-]	CI	38.49	30.07	34.00	23.78	46.68	33.48		40.55	38.69	39.22	37.20	30.49	24.58	35.42	42.55	52.46	35.39
Catchment shape [-]	SHAPE	61.73	40.33	53.36	48.16	70.16	31.24		31.15	19.93	16.60	46.75	46.55	63.42	34.42	40.94	22.03	21.88
Average slope of catchment [%]	AVGSLOPE	44.59	36.76	33.68	37.28	23.48	20.00		24.91	18.98	16.67	18.04	27.87	36.06	17.95	15.96	21.39	16.09
Hypsometric integral [-]	HI	0.48	0.48	0.48	0.50	0.43	0.46		0.46	0.41	0.48	0.46	0.47	0.46	0.47	0.49	0.49	0.48
DEM mean [m]	MDEM	2915	2676	2717	2415	2429	2229		2264	2166	1920	2234	2230	2338	2093	2142	2038	2003
Longest flow path [km]	LFP	97.67	81.58	84.28	43.92	47.64	62.48		24.97	79.96	17.87	36.09	44.74	37.64	42.81	45.87	12.81	26.21
Catchment area [km2]	AREA	1408	1657	1281	531	298	608		736	2072	287	437	588	463	391	325	546	427