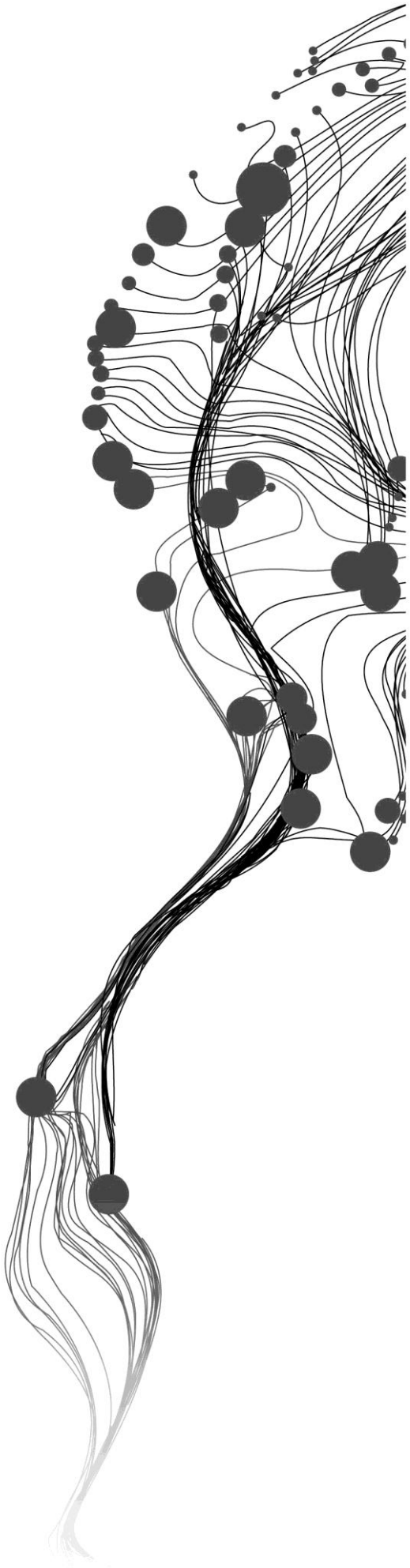
REMOTE SENSING BASED HYDROLOGIC MODELING IN THE BABAHOYO RIVER SUB-BASIN FOR WATER BALANCE ASSESSMENT

MAX EDUARDO SOTOMAYOR MALDONADO
March, 2011

SUPERVISORS:
Dr.ing. T.H.M. Tom Rientjes
Dr. ir. Christiaan van der Tol



REMOTE SENSING BASED HYDROLOGIC MODELING IN THE BABAHOYO RIVER SUB-BASIN FOR WATER BALANCE ASSESSMENT

MAX EDUARDO SOTOMAYOR MALDONADO

Enschede, The Netherlands, March, 2011

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfillment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Water Resources and Environmental Management

SUPERVISORS:
Dr.ing. T.H.M. Tom Rientjes
Dr. ir. Christiaan van der Tol

THESIS ASSESSMENT BOARD:
Prof.Dr. Z. Bob Su (Chair)
Dr. Paolo Regianni (External Examiner, Deltares)

ABSTRACT

The Guayas river system is the largest and probably the most important catchment in Ecuador. Guayas river basin has major problems with water supply for agricultural and industrial production, flooding issues and drought problems in the coastal areas. The information related to climate and hydrology of the Guayas river basin has significant importance to the planning and execution of projects oriented to optimum exploitation of the water resources.

The objective of this study is to simulate the relation between rainfall and runoff of a 3000 km^2 catchment located in the Babahoyo sub-catchment, which is the largest sub-catchment in the Guayas river basin. The aim is to test the applicability of a remote sensing and GIS based distributed rainfall-runoff model for a period of three years (2003-2005).

Hydrometeorological and thematic data were used as model inputs. In addition, TRMM 3B42 Daily product was employed to assess the main driving force of the model, rainfall. MCD15A2 MODIS product was used to account for information regarding the vegetation cover. The LISFLOOD hydrologic model, which runs on PCRaster package, was applied to simulate the catchment runoff in a distributed manner.

Calibration results gave a Root Mean Square Error (RMSE) of 85.85 m^3/s, a Nash-Sutcliffe efficiency (NS) of 0.6 and a Relative Volume Error (RVE) of 46.36%. These values indicate that the simulated hydrograph weakly coincide with the observed hydrograph. The poor network density of the rain gauges in the study area presumably affects the simulation results. A sensitivity analysis of the calibration parameters was performed. UpperZoneTimeConstant (UZTC) influences the recession section of the hydrograph. Small values result in a fast, steep recession, whereas the slope of the falling limb is gentler for higher values. LowerZoneTimeConstant (LZTC) controls much of the baseflow response. Generally, this parameter did not affect largely the obtained simulation. GwPercValue (GPV) controls the baseflow behavior. Higher values results in large amounts of water in the baseflow section of the hydrograph. Increasing b_Xinanjiang (bX) value decreases the infiltration and thus rain rate becomes available for surface runoff; differences of the obtained peaks are relatively small. The PowerPrefFlow (PPF) parameter influences preferential flow of the quickflow component. The resultant hydrographs indicate that the model is not very sensitive to PPF parameter for the catchment of study.

Runoff simulations are performed by rain-gauge data but also the TRMM 3B42 rainfall product is used. This study shows that LISFLOOD model is sensitive to rainfall representation. TRMM rainfall represented by the Inverse Distance Weighted (IDW) interpolation produced different results from those obtained when rain-gauged data was represented by the Thiessen polygons.

Key words: Guayas basin, Babahoyo, TRMM, MODIS, PCRaster, LISFLOOD, rainfall-runoff

ACKNOWLEDGEMENTS

First and foremost I would like to thank God for his mercy all the way through my life.

I owe my deep and most sincere gratitude to my first supervisor Dr. ing. T.H.M. Rientjes for his critical comments, motivation, support and guidance throughout the thesis work. Tom, your continual and excellent supervision has been of great value for me and this research. Furthermore, I am grateful to my second supervisor Dr. ir. Christiaan van der Tol for his constructive comments and his help.

Special and warm thanks are due to Mr. Leonardo Espinoza (SENPLADES) whose logistic support to complete the fieldwork. Leo, I highly appreciate your commitment and cooperation throughout the thesis work. I am also thankful to you for providing me the thematic data for the Guayas river basin.

I would like to thank Ing. Jorge Acosta, Ing. José Luis Rivadeneira, Ing. Alexandra Febres and Egda. Ximena Echeverría (CLIRSEN) whose special support from the initial to the final level enabled me to obtain the ground observations and thematic data of the study area.

I would like to express my appreciation to all my course mates for their being friendly and wonderful events we shared.

I offer my heartily thanks and blessings to all my beloved families for their unwavering support, patience and love throughout my life. Their prayers for me were the main source of inspiration.

Last but not least, I am deeply grateful for all my friends who helped me one way or other to make my stay here in the Netherlands very pleasant and unforgettable.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Water supply in Ecuador is a very serious problem, although the country has an average annual rainfall of 1200 millimeters. The uneven distribution of rainfall and population are the major reason for the water supply problems. Some areas receive only 250 millimeters of rainfall per year while others receive as much as 6000 millimeters per year. Most of the population occupies the mountain regions and the Guayas watershed in the Pacific coastal lowlands where rainfall is relatively low. In contrast, 80 percent of the available water resources in the country are in the sparsely populated Amazon basin. Only 10 percent of the total available water in the country is used, and of this, 97 percent is used for irrigation and 3 percent for domestic and industrial purposes, as it is stated by Encalada (1997).

Figure 1.1 Location of Guayas river basin

Numerous rivers and streams in Ecuador originate on the western slopes of the highlands, end on the Pacific Ocean and drain the coastal zones (see Figure 1.1). The principal drainage systems are the Rio Guayas in the south and the Rio Esmeraldas in the north. The Rio Guayas system is the largest and most important of the region's rivers. From its mouth to the city of Guayaquil, the Rio Guayas is less of a natural river and more of a commercially developed waterway. Upstream from Guayaquil, it divides into the Rio Daule, the Rio Babahoyo, and a multitude of tributaries. These streams enrich the Guayas basin with soils carried down from the Sierra, making the Guayas River basin Ecuador's most fertile agricultural zone (Buckalew J., 1998).

Project 'Generation of geoinformation for land management and valuation of rural lands of the Guayas River Basin' has been launched given that rainfall is relatively low over the Guayas Basin and large amount of natural resources are located in this area. This project is being managed by SENPLADES (Secretaría Nacional de Planificación y Desarrollo), which is a government

institution, and is executed by CLIRSEN (Centro de Levantamientos Integrados por Sensores Remotos). According to CLIRSEN (2009), the project aim to establish the baseline of the natural resources in the study area and to propose use, management and conservation systems in both the headwaters of the basin and micro basins as well as to generate economic alternatives for the rural population.

The module 3 of the project, 'Climate and Hydrology', aims to generate integrated hydro meteorological information for assessing the availability, performance and use of water resources in the Guayas River Basin. This information will be the base to formulate plans for integrated watershed management. The catchment (32000 km^2) is divided into seven sub catchments: Babahoyo River, Daule River, Jujan River, Macul River, Vinces River, Yaguachi River and Drenajes Menores (CLIRSEN Magazine, 2009).

Guayas River basin has major problems with water supply for agricultural and industrial production. It also has flooding problems in the lower valleys and drought problems in the coastal areas. The Guayas basin has one multipurpose project, the Daule-Peripa Dam. The dam provides hydropower, irrigation water, and water supply. It has a 35-megawatt capacity, is 60 to 80 meters high, and contains about 60 million cubic meters of water (Buckalew et al., 1998).

1.2. Relevance of the study

For developing a sustainable water resources management strategy, reliable information on water resources availability must be accessible, for the quantification of the spatial and temporal changes of water balance variables (Wagner et al., 2009). The information related to climate and hydrology of the Guayas river basin has significant importance to the planning and execution of projects oriented to optimum exploitation of the water resources.

Research in this thesis study aims to assess rainfall-runoff relation of a 3000 km^2 catchment located in the Babahoyo sub-basin, which is the largest sub catchment in the Guayas basin. This work was coordinated with the CLIRSEN technical personnel. It serves as pilot analysis and will be the base in order to apply this methodology on the other sub-catchments of the Guayas basin. This study supports the main objective of Module 3 (Climate and Hydrology) of the Guayas Basin project which aims to determine the hydrometeorological performance of the sub-basins and the availability of the water resources.

The rainfall-runoff relationship is strongly dependent on soil, land use and topographic characteristics of the catchment (Jain et al., 2004). Developments in geographic information system (GIS) techniques have enhanced the capabilities to handle large databases describing the variability of land surface characteristics. Remote sensing techniques also can be used to obtain spatial information in digital form on vegetation and rainfall at regular grid intervals with repetitive coverage. Therefore, tools of remote sensing and GIS provide the means of identifying the physical factors that control the process of partitioning of rainfall into runoff and other components. For this reason, it is advisable to define the 3000 km^2 by distributed information.

Advances in digital mapping have provided essential tools to closely represent the 3D nature of a natural landscape. The digital elevation model (DEM) is one of the products of digital mapping technique (Jain et al., 2004). By DEM analysis, topographic variables such as basin geometry, stream networks, slope, flow direction, etc. can be extracted automatically. The integration of a hydrological model with the spatial data handling capabilities of a digital terrain model (DTM) provides information that help to assess, simulate and understand hydrological processes. The grid or cell approach adapts well to the collection of input data on a regular pattern with the use of remote sensing and GIS to represent variation in topographic characteristics within a catchment. The distributed hydrological models that use grid or cell approaches are also capable of quantifying the effect of variability in topographic, meteorological and physiographic characteristics on the catchment runoff (Jain et al., 2004).

Therefore, knowledge of the rainfall-runoff relationship of the catchment of study and spatial distribution of rainfall using satellite observations will help for climate and hydrological studies, water resources planning and management and hydropower development in the region.

1.3. Objectives and research questions

1.3.1. General objective

The general objective of this research is to simulate the relation between rainfall and runoff of a 3000 km^2 catchment located in the Babahoyo sub-basin by using a remote sensing and GIS based distributed hydrological model. This survey serves as a pilot study for integrated water resources management in the Guayas River basin.

1.3.2. Specific objectives

This study focuses on the following specific objectives:

1. To assess the vegetation cover over time through LAI product from MODIS instrument.
2. To determine the potential evaporation and transpiration as important components to the catchment water balance.
3. To compare rainfall amounts between gauge data and TRMM data on a daily base.
4. To assess the sensitivity of the model to rainfall representation in simulating catchment responses.
5. To evaluate whether LISFLOOD model is suitable for applications in catchment systems in The Andes.

1.3.3. Research questions

- What GIS functionalities can be applied to obtain the TRMM grid products on a PCRaster format to be used in LISFLOOD model?
- How the MCD15A2 MODIS product can be processed in order to obtain LAI raster maps?
- How can daily potential evaporation rate from bare soil and daily potential evapotranspiration rate from a reference crop be obtained from daily potential evaporation rate from free water surface on a distributed approach?
- How well the LISFLOOD model simulates the runoff in the catchment of study?
- What are the amount differences between TRMM and gauge rainfall data over the 3000 km^2 catchment?
- What is the influence of rainfall representation on catchment responses?

1.4. Outline of the thesis

The thesis has six chapters. In the first chapter a brief background to the study is given starting from a short description of the situation of the water resources in Ecuador. The importance of the Guayas River Basin is showed and finally the objective of the project 'Generation of geoinformation for land management and valuation of rural lands of the Guayas River Basin' and its Module 3 'Climate and Hydrology' is described. In the same chapter the significance of the study, research objectives and research questions are addressed.

Chapter two presents a literature review and has 5 major parts. First, rainfall-runoff main concepts are described. The second part contains a brief explanation of the software used for running the model. Third, the characteristics of the model itself are presented with an overview of its structure. Then, enlightenment about the TRMM, Terra and Aqua satellites is presented.

A general description of the area where the 3000 km^2 catchment of the study is located, the climate, crops and irrigation is presented in chapter three. In addition, office data, ground observations and remote sensing products used in this study are described in this chapter.

Chapter four discusses the procedures developed in this study to process the remote sensing and the meteorological rainfall data. Additionally, this chapter explains the procedure made to obtain the inputs for the model as topography, land use, soil water characteristics, meteorological and vegetation data to allow for application on a distributed approach. Calibration and its particular objective functions are presented. Moreover, initialization of the model and description of the settings file are explained in this chapter.

The results and discussion sections are presented in chapter five. The chapter discusses the comparison between TRMM 3B42 product and meteorological stations estimates, the output simulation generated by the model, the performed sensitivity analysis, the calibration results and the rainfall representation analysis as a main driving force.

Finally, in chapter six, the conclusions as drawn from this study are presented. Furthermore, some recommendations for future studies are given.

2. LITERATURE REVIEW

2.1. Rainfall – Runoff modeling

Limitations of hydrological measurement techniques and a limited range of measurements in space and time are the main reasons to model the rainfall-runoff processes of hydrology (Beven, 2000). We are not able to measure everything we would like to know about hydrological systems. Consequently, we need a means of extrapolating from available measurements in both space and time, into the future to assess the likely impact of future hydrological change.

Models used in hydrology have equations that involve inputs and state variables. Meteorological inputs are rainfall, potential evapotranspiration, temperature, etc. There are inputs that define the geometry of the catchment that are often considered constant during a simulation. In addition, there are variables that define the time variable boundary conditions. There are the state variables which change during a simulation as a result of the model calculations. There are the initial values of the state variables that define the state of the catchment at the start of a simulation. Finally, there are the model parameters which define the characteristics of the catchment area (Beven, 2000).

The aim of distributed hydrologic modeling is to understand hydrologic catchment behavior and runoff processes by use of cartographic data, satellite data, stream discharge measurements, observations of crops and other vegetation information, meteorology, soil physics, hydrogeology and everything else that is relevant within this context (Abbott et al., 1996).

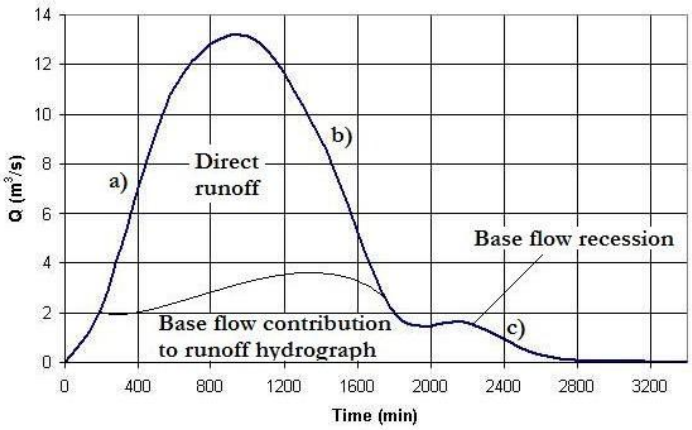


Figure 2.1 Hydrograph sections

The principal output from a rainfall-runoff model is a hydrograph. A hydrograph is interpreted as the integral response function of all upstream processes due to rainfall (Rientjes, 2010). A hydrograph can be divided in three characteristic sections (see Figure 2.1): a) the rising limb, b) the falling limb and c) the baseflow recession. In addition, peakflow section can be added as an element of the direct runoff.

2.2. PCRaster software

The PCRaster package was selected to build a spatially distributed water balance model for the study area. The package is a free raster GIS software that is developed at the Department of Geo-Sciences of the Utrecht University in The Netherlands. According to Van Deursen (1991), several programs have been developed independently to work with PCRaster (i.e. NutShell: a windows shell for easy operation of PCRaster under Windows NT/XP). It also runs under UNIX and MS-DOS.

The software is a set of utilities for hydrological and geomorphological modelling and it is capable to describe changes in the water balance compartments on an hourly, daily and monthly base on each of the cells of the raster system. Furthermore, it counts with a computer language for construction of iterative spatio-temporal environmental models which allows creating models.

The central concept of PCRaster is a discretization of the landscape in space, resulting in cells of information (Van Deursen et al., 1991). Each cell can be regarded as a set of attributes defining its properties, but one which can receive and transmit information to and from neighbouring cells. This representation of the landscape is often referred to as 2.5 D: the lateral directions in a landscape are represented by a set of neighboring cells that make up a map; relations in vertical directions, for instance between soil layers, are implemented using several attributes stored in each cell.

The modules for Cartographic and Dynamic Modelling are integrated with the GIS at a high level, which means that the GIS functions and modelling functions are incorporated in a single GIS and modelling language for performing both GIS and modelling operations.

The dynamic modelling language can be used for building a wide range of models (Van Deursen et al., 1991), from very simple (point) models up to conceptually complicated or physically based models for environmental modelling.

2.3. LISFLOOD model

LISFLOOD is a spatially distributed and physically-based flood simulation model that is capable of simulating the hydrological processes that occur in a catchment. LISFLOOD has been developed by the floods group of the Natural Hazards Project of the Joint Research Centre (JRC) of the European Commission.

According to Van der Knijff et al. (2008), the model is designed to be applied across different spatial and temporal scales. LISFLOOD is grid-based, and applications have used grid cells of 100 up to 5000 meters. Long-term water balance can be simulated (using a daily time step), as well as individual flood events (using hourly time intervals, or even smaller). The model's primary output product is channel discharge and all internal rate and state variables can be written as

output as well. In addition, all output can be written as grids, or time series at user-defined points or areas.

The LISFLOOD (Van der Knijff et al., 2008) model is implemented in the PCRaster (Van Deursen et al., 1991) Environmental Modelling language, wrapped in a Python based interface. The Python wrapper of LISFLOOD enables the user to control the model inputs and outputs and the selection of the model modules. LISFLOOD runs on any operating for which Python and PCRaster are available.

Figure 2.2 gives a summary of the structure of the LISFLOOD model. Basically, the model is made up of the following constituents:

• a 2-layer soil water balance sub-model
• sub-models for the simulation of groundwater and subsurface flow
• a sub-model for the routing of surface runoff to the nearest river channel
• a sub-model for the routing of channel flow

The processes that are simulated by the model include infiltration, interception of rainfall, leaf drainage, evaporation and water uptake by vegetation, surface runoff, preferential flow (bypass of soil layer), exchange of soil moisture between the two soil layers and drainage to the groundwater, sub-surface and groundwater flow, and flow through river channels.

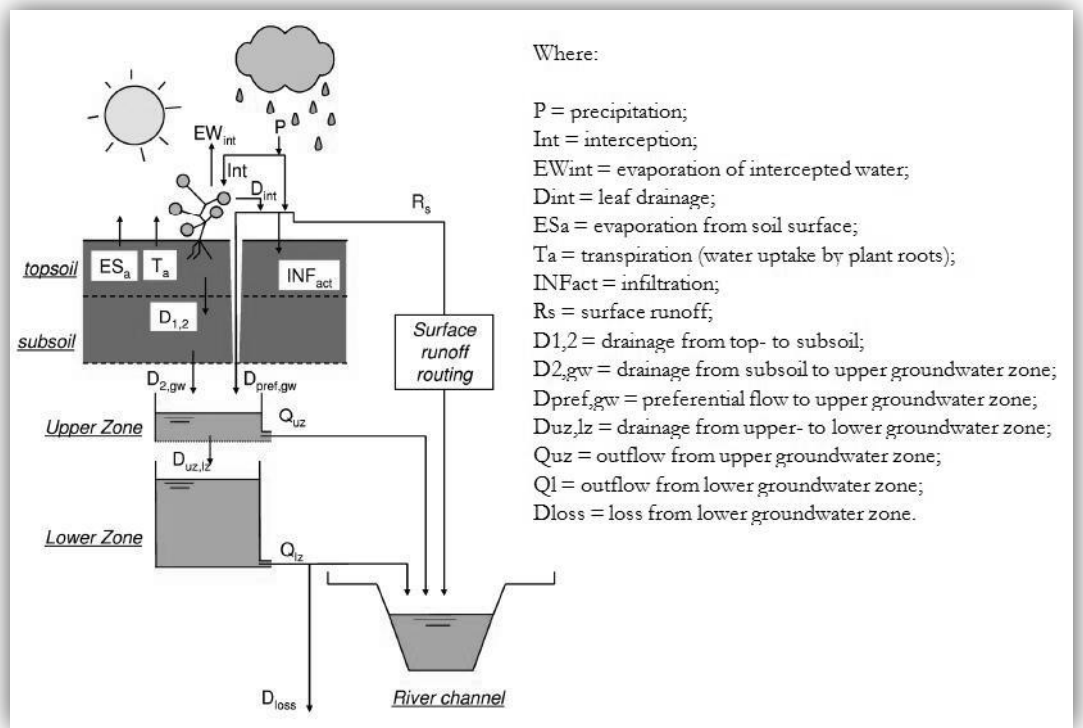


Figure 2.2 LISFLOOD model structure
Source: (Van der Knijff et al., 2008)

A detailed explanation about the processes, equations and assumptions of the LISFLOOD model can be found in the Revised User Manual of LISFLOOD:
http://floods.jrc.ec.europa.eu/files/lisflood/ec_jrc_lisfloodUserManual_JvdK-AdR.pdf.

A crucial step which contributes significantly to the accuracy of the discharge forecasts is the calibration of the LISFLOOD model (Feyen et al., 2007). Owing to the general nature of the LISFLOOD model its application to any given river basin requires that certain parameters of conceptual functions be identified for the particular basin. In the process of calibration, the values of unknown model parameters are tuned such that the model matches the observed predictions as closely as possible. This can be done by manually adjusting the parameters while visually inspecting the agreement between the observed and simulated discharges. However, the subjective and time-consuming nature of the trial-and-error method renders this method unappealing for use on all catchments. In case a large number of catchments for which the model needs to be calibrated, automatic parameter estimation procedure can be the solution. Besides shortening the implementation time this will also enhance the reliability of the calibrated parameters due to a more exhaustive exploration of the parameter space.

LISFLOOD model was especially developed for application in European catchments. According to De Roo et al. (2000), in two pilot transnational European river basins the flooding problem is investigated using the LISFLOOD model: the Meuse catchment, covering parts of France, Belgium and The Netherlands; and the Oder basin, covering parts of The Czech Republic, Poland and Germany.

LISFLOOD is used to obtain the predictions of river discharge in European Flood Alert System (EFAS). This system is developed by Joint Research Centre of the European Commission (Feyen et al., 2007). The system runs on a pre-operational basis using a grid resolution of 5 km for all European river basins larger than 2000 km^2.

2.4. TRMM satellite overview

The Tropical Rainfall Measuring Mission (TRMM) is a joint endeavor between NASA and Japan's National Space Development Agency. It is designed to monitor and to study tropical rainfall and the associated release of energy that helps to power the global atmospheric circulation, shaping both global weather and climate (NASA, 2010).

Before TRMM's launch measurements of the global distribution of rainfall at the Earth's surface had uncertainties of the order of 50% and the global distribution of vertical profiles of precipitation was far less well determined (NASA, 2010). TRMM is providing spaceborne rain radar and microwave radiometric data that measures the vertical distribution of precipitation over the tropics in a band between 35 degrees north and south latitudes. Such information enhances the understanding of the interactions between the sea, air and land masses which produce changes in global rainfall and climate. TRMM observations also help improve modelling of

tropical rainfall processes and their influence on global circulation leading to better predictions of rainfall and its variability at various time scales.

In its rainfall measurement package, TRMM has the following instruments: Precipitation Radar (PR), the TRMM Microwave Imager (TMI), and the Visible and Infrared Radiometer System (VIRS).

2.4.1. Precipitation Radar (PR)

According to (NASA, 2010), the Precipitation Radar (PR) has a horizontal resolution at the ground of about five kilometres and a swath width of 247 kilometres. One of its most important features is its ability to provide vertical profiles of the rain and snow from the earth surface up to a height of about 20 kilometres. The Precipitation Radar is able to detect fairly light rain rates down to about 0.7 millimetres per hour. At intense rain rates, where the attenuation effects can be strong, new methods of data processing have been developed that help correct for this effect. The Precipitation Radar is able to separate out rain responses for vertical sample sizes of about 250 meters when looking straight down. It carries out all these measurements while using only 224 watts of electric power—the power of just a few household light bulbs. Although weather radars on the ground have been used ever since World War II to estimate rainfall, there were many technical challenges that had to be overcome before an instrument of this kind could be used from space.

2.4.2. TRMM Microwave Imager (TMI)

The Tropical Rainfall Measuring Mission's (TRMM) Microwave Imager (TMI) is a passive microwave sensor designed to provide quantitative rainfall information over a wide swath under the TRMM satellite. By measuring the minute amounts of microwave energy emitted by the Earth and its atmosphere, TMI is able to quantify the water vapour, the cloud water, and the rainfall intensity in the atmosphere. It is a relatively small instrument that consumes little power (NASA, 2010).

TMI is based on the design of Special Sensor Microwave/Imager (SSM/I) which has been flying continuously on Defense Meteorological Satellites since 1987. The TMI measures the intensity of radiation at five separate frequencies: 10.7, 19.4, 21.3, 37, 85.5 GHz. These frequencies are similar to those of the SSM/I, except that TMI has the additional 10.7 GHz channel designed to provide a more-linear response for the high rainfall rates common in tropical rainfall. The other main improvement of TMI is due to the improved ground resolution. This improvement, however, is not the result of any instrument improvements, but rather a function of the lower altitude of TRMM (402 kilometres compared to 860 kilometres of SSM/I). TMI has an 878-kilometer wide swath on the surface. The higher resolution of TMI on TRMM, as well as the additional 10.7 GHz frequency, makes TMI a better instrument than its predecessors. The additional information supplied by the Precipitation Radar further helps to improve algorithms (NASA, 2010).

2.4.3. Visible and Infrared Scanner (VIRS)

Visible and Infrared Scanner (VIRS), as its name implies, senses radiation coming up from the Earth in five spectral regions, ranging from visible to infrared, or 0.63 to 12 micrometers. VIRS is included in the primary instrument package for two reasons. First is its ability to delineate rainfall. The second is to serve as a transfer standard to other measurements that are made routinely using POES and GOES satellites. The intensity of the radiation in the various spectral regions (or bands) can be used to determine the brightness (visible and near infrared) or temperature (infrared) of the source.

A variety of techniques use the Infrared (IR) images to estimate precipitation. Higher cloud tops are positively correlated with precipitation for convective clouds (generally thunderstorms) which dominate tropical (and therefore global) precipitation accumulations. One notable exception to this rule of thumb is the high cirrus clouds that generally flow out of thunderstorms. These cirrus clouds are high and therefore "cold" in the infrared observations but they do not rain. To differentiate these cirrus clouds from water clouds (cumulonimbus), a technique which involves comparing the two infrared channels at 10.8 and 12.0 micrometers is employed. Nonetheless, IR techniques usually have significant errors for instantaneous rainfall estimates. VIRS uses a rotating mirror to scan across the track of the TRMM observatory, thus sweeping out a region 833 kilometres wide as the observatory proceeds along its orbit. Looking straight down (nadir), VIRS can pick out individual cloud features as small as 2.4 kilometres (NASA, 2010).

2.5. Terra and Aqua spacecrafts

2.5.1. Aqua spacecraft

Aqua, Latin for water, is a NASA Earth Science satellite mission named for the large amount of information that the mission will be collecting about the Earth's water cycle, including evaporation from the oceans, water vapor in the atmosphere, clouds, precipitation, soil moisture, sea ice, land ice, and snow cover on the land and ice (NASA, 2011). Additional variables also being measured by Aqua include radiative energy fluxes, aerosols, vegetation cover on the land, phytoplankton and dissolved organic matter in the oceans, and air, land, and water temperatures. The Aqua mission is a part of the NASA-centered international Earth Observing System (EOS). Aqua was formerly named EOS PM, signifying its afternoon equatorial crossing time (NASA, 2011).

Aqua was launched on May 4, 2002, and has six Earth-observing instruments on board, collecting a variety of global data sets. Aqua was the first member launched of a group of satellites termed the Afternoon Constellation (NASA, 2011).

2.5.2. Terra spacecraft

Terra is a multi-national, multi-disciplinary mission involving partnerships with the aerospace agencies of Canada and Japan. Managed by NASA's Goddard Space Flight Center, the mission also receives key contributions from the Jet Propulsion Laboratory and Langley Research Center (NASA, 2011).

On December 18, 1999, NASA launched Terra, the first of a series of large satellites meant to monitor the health of our planet. Terra carries five instruments, including two from Japan and Canada, that together track Earth's land, atmosphere, and ocean (NASA, 2011). Terra's primary mission is to answer the question: How is the Earth changing and what are the consequences of change for life on Earth?

Terra has observed changes all over the world due to irrigation projects, deforestation, artificial islands projects, climate and environmental change, etc (NASA, 2011).

3. STUDY AREA AND MATERIALS

3.1. Study area description

3.1.1. Geographical information

The study area comprehends a catchment which has an area of 2910 km^2. The region has hilly and plane topography and the highest altitude in the basin is 4474 m.a.s.l. The catchment is located between 667350 E, 9841640 N and 746129 E, 9897295 N. These coordinates are on the Projected Coordinate System WGS-1984 UTM Zone 17S (Figure 3.1).

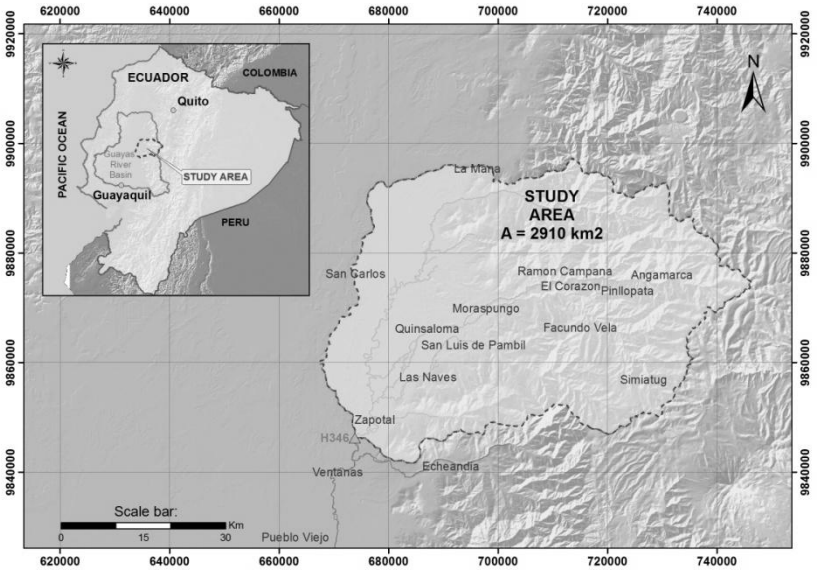


Figure 3.1 Location of the study area

This catchment belongs to the Babahoyo River basin, in western Ecuador, which is fed by tributaries that originate in the Andes Mountains. It joins the Daule River to form the Guayas River, which discharges to the Pacific Ocean.

3.1.2. Historical information

As indicated by CLIRSEN (2009), in recent years the effects and impacts by the intensive use of the natural resources over the Guayas basin are heading to depletion, destruction and degradation of these. It creates disequilibrium of the ecosystems and affects the ecological integrity.

Furthermore, we can observe an accelerated transformation of the territories, landscapes, ecological processes, erosion and soil degradation over the Guayas basin. These are related to expanding the agricultural frontier, logging and deforestation, inappropriate agro-production practices, expansion of urban boundaries, overgrazing, and use of agrochemicals and other sources of pressure that can be observed in the area.

Consequently, it is essential to develop adequate management strategies of political, social, economic and environmental conditions to recover the productive capacity of the people who live in the rural area of the Guayas basin.

The project 'Generation of geoinformation for land management and valuation of rural lands of the Guayas River Basin' will derive basic information as land use, land cover, production systems, landforms, soils, economic situation and hydrology of the catchment.

3.1.3. Climate

The upper catchment limit of this study passes through The Andes mountains (see Figure 1.1), and this aspect largely influences on the rainfall over the area. Evaporation from ocean surfaces is the chief source of moisture for precipitation (Linsley et al., 1982). However, nearness to the oceans does not necessarily lead to adequate precipitation, as evidenced by many desert islands. The location of a region with respect to the general circulation, latitude, and distance to a moisture source is primarily responsible for its climate. Orographic barriers often exert more influence on the climate of a region than nearness to a moisture source does.

According to Peñaherrera (2009), annual rainfall varies from 1200 to 2600 mm while potential evapotranspiration varies from 1100 to 1550 mm over the Babahoyo area. The daily mean temperature varies from 22°C to 27°C. The average wind speed is 0.8 m/s and the mean relative humidity is larger than 85%.

Water deficit for agricultural activities goes from 250 to 550 mm with a dry period that ranges between 110 and 70 days from June to December and a growing season of 140 - 210 days between December and June (Peñaherrera, 2009).

3.1.4. Land cover

Babahoyo sub-basin is covered mostly by rice, soy and banana crops. Other cultivated crops are grass, cocoa, corn and estate gardens. A small part of the agricultural areas are covered by sugar cane for industrial use, teak, oil palm, pineapple, fern tree, passion fruit, balsa, sugar cane for craft use, orange and snuff cover. There is also natural coverage as moist shrubs and natural grass. The least part is covered by cities, towns and infrastructure. It is a high productive zone although this region has vast issues of floods due to its topography (Almeida et al., 2009).

3.1.5. Irrigation

An important part of the region is irrigated by the use of irrigation channels of CEDEGE (Comisión de Estudios para el Desarrollo de la Cuenca del Río Guayas) which was a public commission in charge of making necessary research and studies for the development of the River Guayas Basin. The most common type of irrigation over the area is by ways alongside the channels. The most important irrigation project in the study area is the Babahoyo Irrigation Project. This project was managed by CEDEGE and involved the construction of inland

waterways for irrigation and draining, also pumping canals, paths and protection dikes against floods.

3.2. Materials

3.2.1. Office collected data

Since 2009 CLIRSEN has been working on processing the hydro-meteorological data. Moreover, they have collected a variety of thematic information about the Guayas River basin. For this reason, it was essential to collect all of these data from SENPLADES and CLIRSEN offices which are located in Quito – Ecuador. This information was evaluated and standardized in order to work in the same coordinate system and units. The office data collected is showed in Figure 3.4 and described below.

a) A digital Land Use Map of Ecuador, scale 1:250.000, from MAG (Ministerio de Agricultura y Ganadería) and CLIRSEN.
b) A digital Soils Map of Ecuador, scale 1:200.000, produced by DINAREN (Dirección Nacional de Recursos Naturales).
c) 90 meters DEM extracted from SRTM by CLIRSEN.
d) Ground observations as daily potential evaporation pan, daily rainfall given by INAMHI (Instituto Nacional de Meteorología e Hidrología) and daily discharges for the H346 station called "Zapotal en Lechugal", from CLIRSEN.

The density of the meteorological stations of this study is approximately one station per 340 km^2. According to Linsley et al. (1982) the minimum densities of precipitation networks have been suggested for general hydrometeorological purposes by the World Meteorological Organization (WMO). For mountain regions of temperate, mediterranean, and tropical zones, one station per 100 to 250 km^2 is recommended. Table 3.1 list the meteorological stations with rainfall data available.

Table 3.1 Meteorological stations

CODE	NAME	PERIOD OF RECORD	
M006	PICHILINGUE	1986	2006
M123	EL CORAZON	1986	2006
M124	SAN JUAN LA MANA	1989	2006
M368	MORASPUNGO-COTOPAXI	1995	2006
M470	MOCACHE	1986	2006

Stations M006 and M123 have evaporation pan records available. In addition, the H346 hydrometric station has 22 years with daily discharge data for the period 1984-2005.

Figure 3.2 shows a visual impression of the distribution of the observed daily rainfall for the period 2003-2005. The vast majority of the values are in the range of 0-0.5 mm.

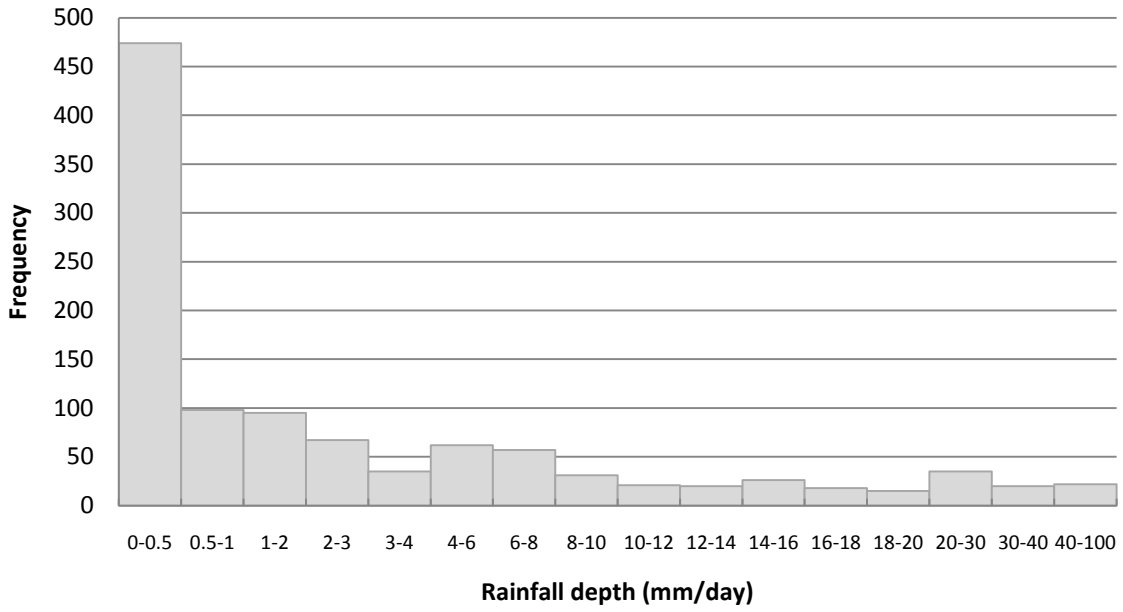


Figure 3.2 Daily rainfall histogram (2003-2005) based on observed rainfall

The relationship between the observed discharge and the average rainfall obtained from the rain gauges were compared by means of the double-mass curve. Double-mass curve tests the reliability of the observed discharge by comparing its accumulated values with the concurrent accumulated values of mean rainfall for a group of meteorological stations.

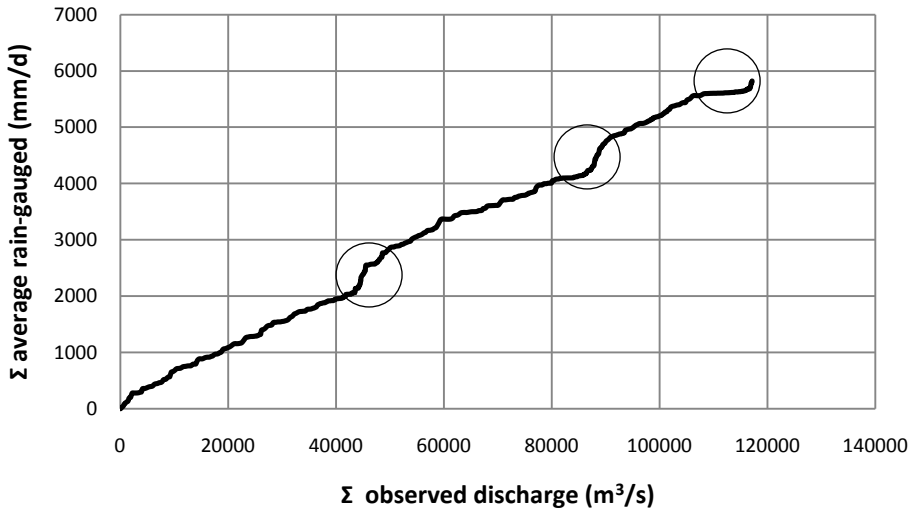


Figure 3.3 Double-mass curve. Observed discharge vs. rain-gauged

Figure 3.3 shows three slope changes that indicate a doubtful relation between the average rainfall and the observed discharge.

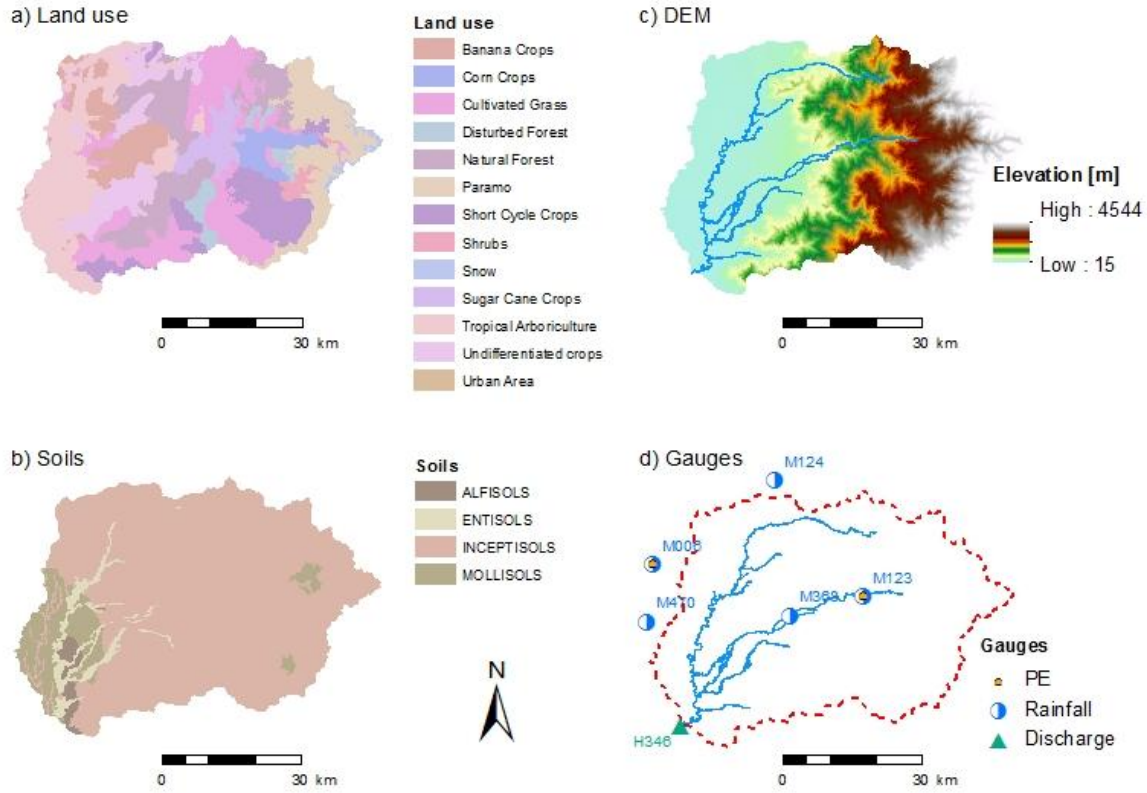


Figure 3.4 Office collected data

The average rainfall rate obtained from the meteorological stations (M006, M123, M124, M368 and M470) and the hydrograph obtained from the discharge station (H346) are shown in Figure 3.5.

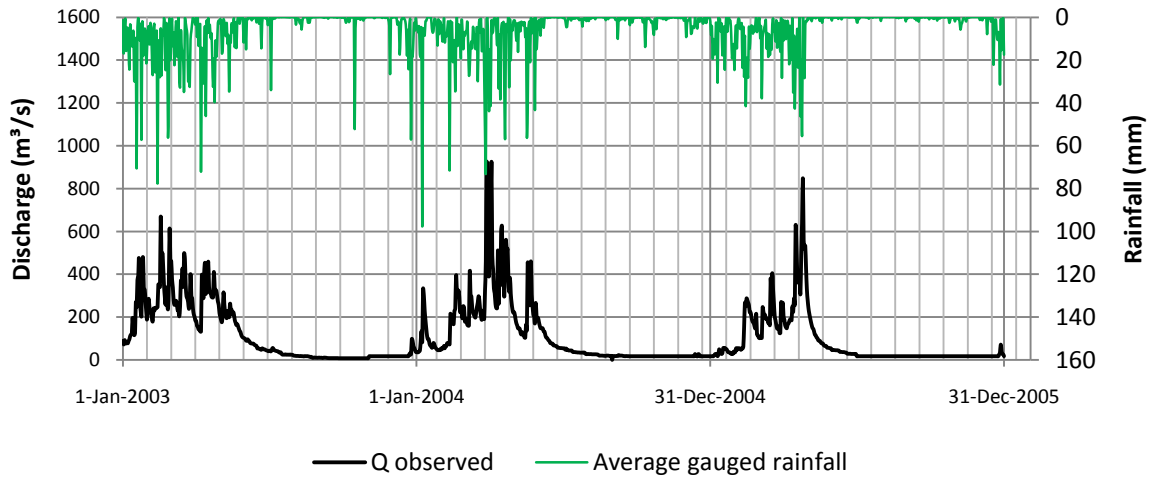


Figure 3.5 Observed discharge vs. average gauged rainfall

Figure 3.5 moreover shows that the observed discharge do not respond well to all of the rainfall events, especially during the dry season and at the beginning of the year 2005.

3.2.2. Satellite data products

In this study two kinds of satellite products have been used: *Daily TRMM and Others Rainfall Estimate (3B42 V6 derived)* from TRMM (Tropical Rainfall Measuring Mission) and *Leaf Area Index* from MODIS (Moderate Resolution Imaging Spectroradiometer).

TRMM 3B42 (daily) algorithm

The purpose of the 3B42 algorithm is to produce TRMM-adjusted merged-infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates (NASA, 2010). The algorithm consists of two separate steps. The first step uses the TRMM VIRS and TMI orbit data (TRMM products 1B01 and 2A12) and the monthly TMI/TRMM Combined Instrument (TCI) calibration parameters (from TRMM product 3B31) to produce monthly IR calibration parameters. The second step uses these derived monthly IR calibration parameters to adjust the merged-IR precipitation data, which consists of GMS, GOES-E, GOES-W, Meteosat-7, Meteosat-5, and NOAA-12 data. The final gridded, adjusted merged-IR precipitation (mm/hr) and RMS precipitation-error estimates have a 3-hourly temporal resolution and a 0.25-degree by 0.25-degree spatial resolution. Spatial coverage extends from 50 degrees south to 50 degrees north latitude.

The daily accumulated rainfall product is derived from this 3-hourly product. The data are stored in flat binary. The file size is about 2.25 MB (uncompressed). The data used in this effort were acquired as part of the activities of NASA's Science Mission Directorate, and are archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (http://disc.sci.gsfc.nasa.gov).

More details of the TRMM 3B42 algorithm are explained on http://trmm.gsfc.nasa.gov/3b42.html and the file specifications are available at http://pps.gsfc.nasa.gov/tsdis/Documents/ICSVol4.pdf.

MCD15A2 MODIS product

The MODIS instrument is operating on the Terra and Aqua spacecrafts. It has swath width of 2,330 km and views the entire surface of the Earth (USGS). Its detectors measure 36 spectral bands and it acquires data at spatial resolutions of 250-m, 500-m, and 1,000-m. The MODIS main instrument specifications are shown in the table below.

Table 3.2 MODIS instrument specifications

Orbit	705 km, 10:30 a.m. descending node (Terra) or 1:30 p.m. ascending node (Aqua), sun-synchronous, near-polar, circular
Scan Rate	20.3 rpm, cross track
Swath Dimensions	2330 km (cross track) by 10 km (along track at nadir)

Telescope	17.78 cm diam. off-axis, afocal (collimated), with intermediate field stop
Size	1.0 x 1.6 x 1.0 m
Weight	228.7 kg
Power	162.5 W (single orbit average)
Data Rate	10.6 Mbps (peak daytime); 6.1 Mbps (orbital average)
Quantization	12 bits
Spatial Resolution	250 m (bands 1-2) 500 m (bands 3-7) 1000 m (bands 8-36)
Design Life	6 years

Source: http://modis.gsfc.nasa.gov

According to US Department of the Interior (2010), the level-4 MODIS global Leaf Area Index (LAI) and the Fraction of Photosynthetically Active Radiation (FPAR) product is composited every 8 days at 1-kilometer resolution on a Sinusoidal grid. The MCD15A2 product includes the following Science Data Sets: LAI, FPAR, and a set of quality rating, and standard deviation layers for each variable. Table 3.3 shows the MCD15A2 data set characteristics. These data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (lpdaac.usgs.gov).

The algorithm uses multispectral surface reflectances and a land cover classification map as input data to retrieve global LAI and FPAR fields. The objectives are to evaluate its performance as a function of spatial resolution and uncertainties in surface reflectances and the land cover map. The algorithm can retrieve a value of LAI/FPAR from the reflectance data and justifies the use of more complex algorithms, instead of NDVI-based methods. The algorithm was tested to investigate the effects of vegetation misclassification on LAI/FPAR retrievals (NASA, 2010).

Table 3.3 MCD15A2 product characteristics from MODIS

Characteristic	**Description**
Temporal Coverage	March 5, 2000 – present
Area	~10° km x 10° Lat/Lon
Image Dimensions	1200 rows x 1200 columns
Spatial Resolution	1 km
File Size	2 MB
Projection	Sinusoidal
Data Format	HDF-EOS
No. of Science Data Sets (SDS)	6

Source: https://lpdaac.usgs.gov/lpdaac

MCD15A2 MODIS product was obtained from https://lpdaac.usgs.gov/lpdaac/get_data/data_pool.

4. METHODS

4.1. TRMM rainfall data processing

TRMM 3B42 daily data was obtained on a NetCDF gridded format from January 1st 2003 to December 31st 2005. These 1096 files (note that there is one more day because 2004 is a Leap Year) were downloaded by the 'Convert to NetCDF' service of the web site (http://disc.sci.gsfc.nasa.gov). It creates a URL List which was saved to a specific location on the local computer as "myfile.dat". Then, the data were obtained using 'wget' software by typing on the command line: "`wget --content-disposition -i myfile.dat`". "myfile.dat" should be saved on the same directory of 'wget' software. Figure 4.1 shows a raster NetCDF from TRMM 3B42 daily product before processing.

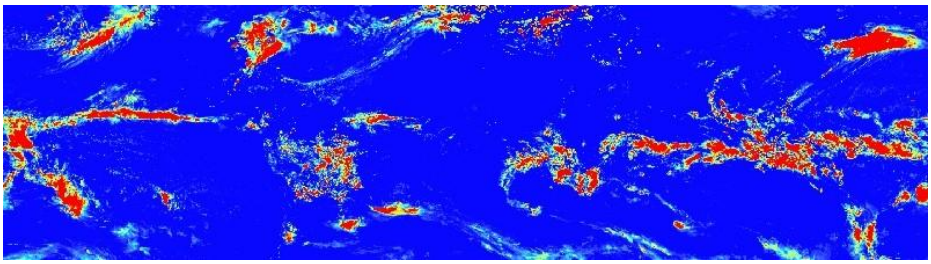


Figure 4.1 Raster NetCDF of TRMM 3B42, January 1st 2003

In order to obtain the rainfall maps for each daily time step, TRMM 3B42 NetCDF format files were processed and converted to a text format files. To make this process automatically for all the files, as part of this study a specific routine was developed using 'Model Builder' from ArcGIS (see Figure 4.2).

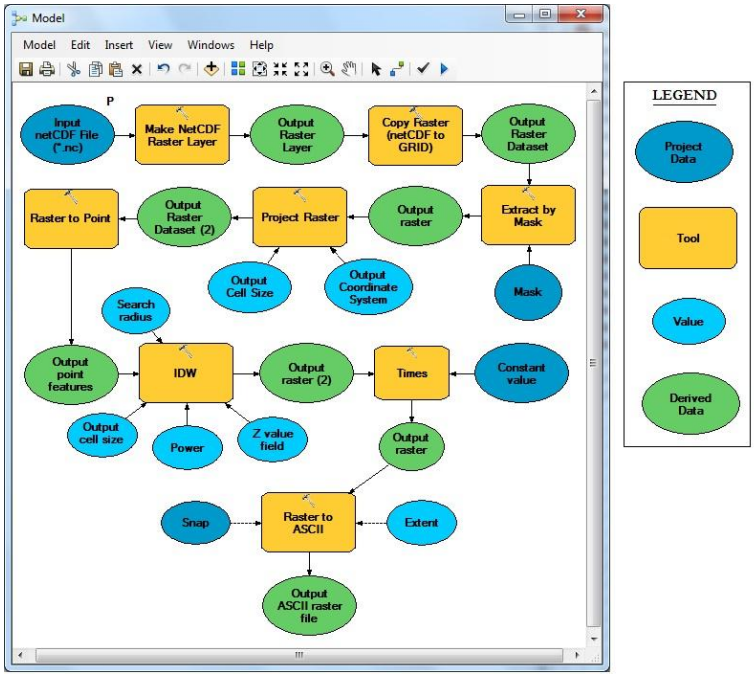


Figure 4.2 TRMM 3B42 processing routine

The first step of this routine is to read the NetCDF file into ArcGIS. For this tool, the input is set as a parameter ('P') which means that the user has to enter this file. ArcGIS allows entering multiple files at once by running the model as a batch. The output of this tool is imported (copied) to a GRID (ArcGIS) format since the software cannot process NetCDF files directly. Once the file has GRID format, it is extracted by the study area polygon (Mask). Note that this polygon is larger than the real area given that pixels out of the catchment are needed for interpolation.

After that, the output raster is projected to 'WGS-1984 UTM Zone 17S' coordinate system. The output coordinate system and the output pixel size value of 25000 meters have to be entered. The routine saves this information and keeps the values for each run.

The next step is to convert from raster to point feature. This is needed to obtain the rainfall values at the center of each pixel. Then, these points are interpolated using an inverse distance weighted (IDW) technique. After various tests, IDW interpolation was performed with the following variables: search radius (8), output cell size (1000), Power (2) and the Z value field (GRID_CODE). Search radius defines which surrounding points are used to control the raster. It was set to eight (number of nearest input sample points to be used to perform interpolation) to reduce boundary effects. The cell size (at which the output raster will be created) was set as 1000 meters to match the LAI raster spatial resolution. The Power controls the significance of surrounding points on the interpolated value. A higher Power results in less influence from distant points. This parameter was tested with values '1' and '2' where '2' showed that rainfall was distributed in a plausible manner (see Figure 4.3). Finally, 'GRID_CODE' is the field that holds the rainfall value for each point.

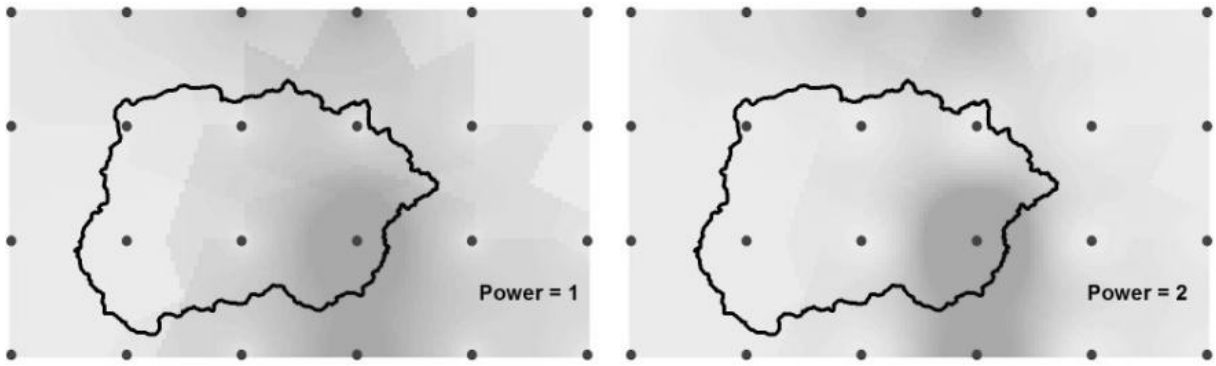


Figure 4.3 IDW Power difference representation

Then, the output raster files are multiplied by the constant value of 3 since each of the layers is containing the precipitation in mm/hr for a period of three hours.

Lastly, the raster files are converted to ASCII files (*.txt) since PCRaster creates the time series maps for modeling from this format. Two additional inputs are required by this tool, the spatial extent and the snap raster. This information is given by the raster file which defines the spatial domain. The characteristics of this raster file is explained in section 4.5.1 Input files.

4.2. Rain gauges data procedure

Rainfall data are taken from five meteorological stations: M006, M123, M124, M368 and M470 (see Section 3). Time series are distributed over the study area by converting the rain gauges location points to an output feature class of Thiessen polygons (see Figure 4.4). Thiessen

polygons have the unique property that each polygon contains only one input point and any location within a polygon is closer to its associated point than to the point of any other polygon.

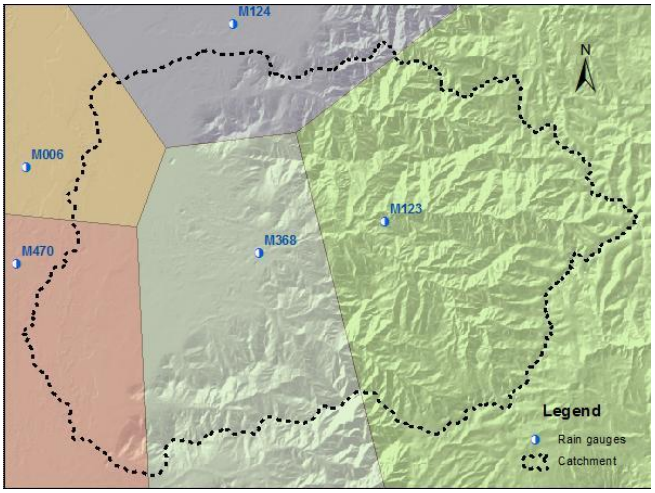


Figure 4.4 Rainfall Thiessen polygons

Rainfall time series data are organized in a table. Rows correspond to the meteorological station codes and columns correspond to the respective day. An example is shown in the table below:

Table 4.1 Rainfall data from meteorological stations

COD	1	2	3	4	5	6…
M006	0.9	32.7	9.3	0	0.4	0.7…
M123	0.50	6.30	3.40	1.40	9.20	11.80…
M124	3	35.2	3.2	0	9.3	9.2…
M368	1.6	4.4	9.2	3.6	31.9	3…
M470	1.2	6.3	0	0	28.2	0…

Subsequently, this table of time series was joined to the Thiessen feature class. Then, this polygon feature was converted to raster dataset on a batch mode based on each day (field). As a result, 1096 rainfall raster maps representing rainfall from January 1st 2003 to December 31st 2005. Finally, these files were converted to ASCII files.

4.3. MODIS Leaf Area Index (LAI) procedure

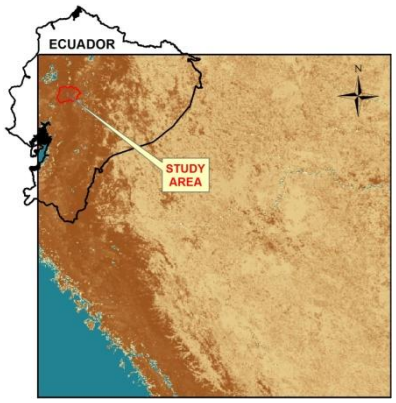


Figure 4.5 MODIS image for January 1st, 2003

Leaf Area Index data is an important input of the model since LAI is used in the equations that describe several processes of the LISFLOOD model as interception, evaporation of intercepted water, water uptake by plants roots and transpiration, and direct evaporation from the soil surface.

Leaf Area Index was obtained from the MCD15A2 MODIS product, which belongs to the MCD15A2.5 data set. This is enclosed in the MOTA (MODIS Land collections/granules combined from the Terra and Aqua missions) data group. A 'data group' is a grouping data collection by instrument, mission, and major discipline. Those are the criteria for searching the LAI images. Figure 4.5 shows the raw MODIS image for January 1st, 2003.

Leaf Area Index (LAI) is a variable that changes relatively slowly over time. For that reason, only one image per month from January 2003 to December 2005 was selected to represent vegetation cover. Temporal resolution of MODIS instrument is 8 days. Therefore, several images per month are available. The image with less cloud coverage percentage of each month was selected. Then, the images were downloaded.

The LAI products are in HDF-EOS format (see Table 3.3) and ERDAS IMAGINE was used to process the images. The first step was importing the images to ERDAS format (*.img). In this step, the 'Lai_1km' field was selected. Then, the images were reprojected based on the following settings: output projection (UTM WGS84 South, Zone 17), units (meters), output cell size (1000 x 1000 meters) and resample method (Nearest Neighbor). This procedure allows projecting, subset and resampling the images in one single step. Furthermore, the images were processed as a batch to speed up the data processing. Resampling was required since the original pixel size is not exactly 1000 x 1000 meters but 1049.52 x 936.54 meters. When subsetting the images, the respective coordinates were modified since Erdas takes the center point of the pixel as a reference while PCRaster uses the left lower corner of the pixel.

Afterward, images were converted from digital numbers (DN) to LAI values by multiplying by 0.1 using 'Times' tool of ArcGIS as a batch mode. Finally, the raster datasets were exported to ASCII files (*.txt).

4.4. Potential evapotranspiration data process

Evaporation and water uptake and subsequent transpiration by vegetation are important components of the water balance (Van der Knijff, 2008). The simulation of these processes in LISFLOOD involves three different climatic variables:

1. Potential evaporation of open water surface (EWO)
2. Potential reference evapotranspiration (ETO)
3. Potential soil evaporation (ESO)

There are different methods to estimate evapotranspiration. Remote sensing is a valuable tool that can be used on energy balance models to estimate evapotranspiration. Energy balance models are developed to estimate atmospheric turbulent fluxes and surface evaporative fraction at the time of the image using satellite earth observation data. However, cloud free images must be collected in order to obtain more accurate results (Gonzales, 2010). Given the difficulty of

finding cloud free images for the study area, time and meteorological data constraints, potential evapotranspiration was obtained by the procedure explained below. This procedure is part of the LISFLOOD model approach as presented in Van der Knijff et al. (2008).

4.4.1. Potential evaporation of open water surface (EWO)

Natural evaporation can be measured either as the rate of loss of liquid water from the surface or as the rate of gain of water vapor by the atmosphere (Maidment, 1992). Measurements in the liquid phase assume a closed system, such as an evaporation pan, and deduce evaporation as the net loss of water from that closed system. Because of its apparent simplicity, the evaporation pan is probably the instrument used most widely to estimate potential evaporation.

There are two meteorological stations with evaporation pan data available in the catchment; M006 and M123 (see Figure 4.6). Therefore, the study area was divided in two areas that are represented by each of these meteorological stations respectively. To materialize this division, the DEM was reclassified in such way that the upper pixels (above 500 m.a.s.l.) have value of 123 and lower pixels (below 500 m.a.s.l.) have the value of 6 (these values act as station codes). By this procedure, data from station M123 describe potential evaporation of the mountainous area while M006 characterize potential evaporation in the plane area of the catchment. Then, this raster dataset was converted to a polygon feature that is shown in the figure below.

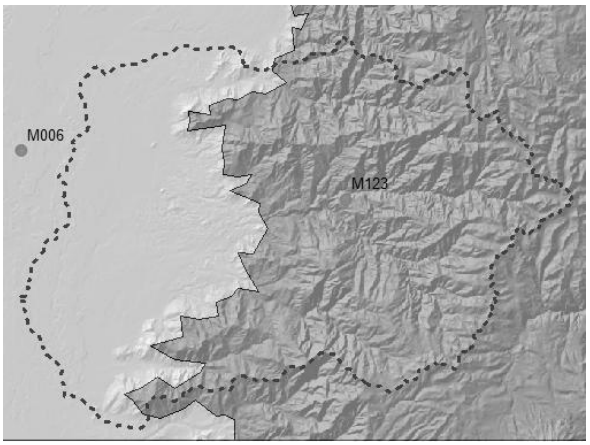


Figure 4.6 Area division for potential evaporation

On the other hand, the potential evaporation values obtained from the meteorological stations were organized in a table of two rows (one for values of M006 and the other for values of M123). The columns of this table are the days, from 1 (January 1st 2003) to 1096 (December 31st 2005). An example of this table with the first 7 days of the time series is showed on Table 4.2.

Table 4.2 Potential evaporation values

Cod	1	2	3	4	5	6	7…
6	2.2	0	0	3.7	1.3	3.1	1.9…
123	1.2	0.7	2.4	1.3	0.5	2.9	2.7…

The created polygon feature was 'joined' to this table and exported (copied) in order to have a polygon feature with its own attribute table. The next step was converting this polygon feature to raster on a batch mode based on the columns of the attribute table that represent the day

number. This is done with the purpose of having one potential evaporation raster map for each time step. Finally, this raster datasets were converted to ASCII files (*.txt) to be read by PCRaster. All of this procedure was developed on ArcGIS software.

4.4.2. Potential reference evapotranspiration rate (ETO)

The evaporation pan can differ significantly from the potential evaporation of surrounding vegetation. Consequently, it is necessary to fix these differences using empirical 'pan coefficients' (Maidment, 1992). Since pan data are widely available and much used for estimating crop water use for irrigation purposes, empirical pan coefficients have been derived. Maidment (1992) has suggested values for the Pan Coefficient (K_{pan}), which relates potential reference evapotranspiration (ETO) to measured pan evaporation (EWO) in the equation ETO = K_{pan} EWO.

For this study, the K_{pan} value of '0.85' was selected considering: a pan surrounded by short green crop, mean relative humidity >70% and wind speed <1 m/s. These characteristics match the study area behavior that presents wet vegetation, average wind speed of 0.8 m/s and mean relative humidity larger that 85% (see section 3.1.3 Climate).
Consequently, potential reference evapotranspiration raster maps were obtained by multiplying the potential evaporation of open water surface raster maps by '0.85' using the 'Times' tool of ArcGIS as a batch mode to end with the conversion from raster to ASCII.

4.4.3. Potential soil evaporation rate (ESO)

The Food and Agriculture Organization Irrigation and Drainage (FAO) calculates daily evaporation from soil by ESO = K_e ETO (Allen et al., 2005). The daily evaporation coefficient is denoted as K_e and the general equation for K_e is: $K_e = (K_{c\ max} - K_{cb})$.

$K_{c\ max}$ represents the higher crop coefficient value for the surface under full vegetation and completely wet surface and K_{cb} is the basal crop coefficient. Because the initial period for annual crops is defined as when there is less than 10% vegetation, K_{cb} during the initial period for annual crops represents a condition of bare, dry soil. Because the objective of the crop coefficient during the initial period ($K_{c\ ini}$) calculation is to predict total evaporation, the value for K_{cb} is set to zero. Thus, $K_e = K_{c\ max}$ and ESO = $K_{c\ max}$ ETO.

During periods having vegetation, $K_{c\ max}$ is predicted as a function of crop height, wind speed, and relative humidity. However, during the initial period, where crop height is zero or nearly zero, a constant value for $K_{c\ max} = 1.15$ is typically used. This value is consistent with Fig. 6 of FAO-24 and with findings from other studies (Allen et al., 2005).

To support this concept, an additional theory from FAO-56 (FAO, 1990) is explained. In order to predict $K_{c\ ini}$ under variety of soil and wetting conditions, FAO-56 has presented a function that was used to develop Figure 4.7. (Allen et al., 2005). This figure represent two wetting conditions (small infiltration depths and large infiltration depths) and two soil texture class (fine and medium textures). The advantage of multiple figures for $K_{c\ ini}$ is the improvement in predicting total water consumption early in crop development periods under varying conditions of wetting and soil texture. Nevertheless, the figure shows that the maximum value of $K_{c\ ini}$ is approximately '1.15'.

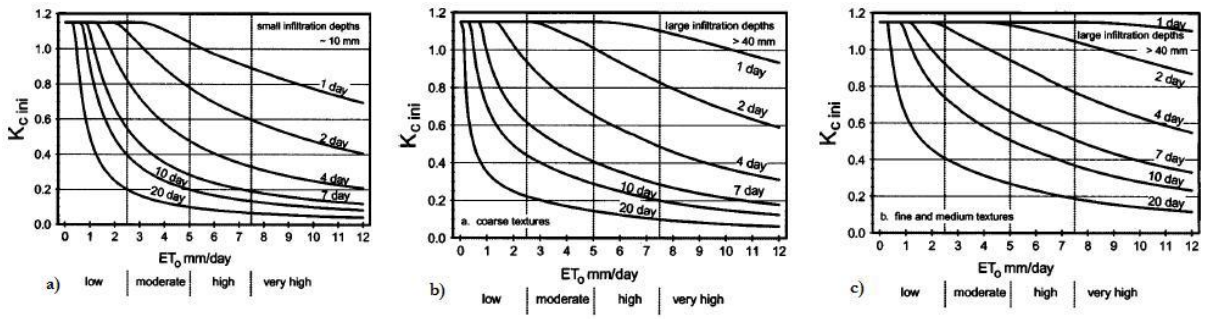


Figure 4.7 $K_{c\ ini}$ related to the level of ETO

With this background, the potential soil evaporation rate (ESO) maps were calculated by multiplying the potential reference evapotranspiration rate (ETO) maps by the constant value of '1.15'. In order to perform this raster operation, 'Times' tool from ArcGIS was applied. Then, these raster datasets were converted to ASCII files.

4.5. Model setup

Model setup involves the collection of the input files (maps and tables), preparing the settings file (constants and parameters) and the initialization. In addition, the outputs of the model are presented.

4.5.1. Input files

All model inputs have been provided as maps and tables. The maps are grid files in PCRaster format and tables are text files readable by any basic text editing program. As explained above, the yield maps from previous sections (3, 4.1, 4.3 and 4.4) were converted to ASCII files. These files were converted to PCRaster maps using the PCRaster application 'asc2map'. This application needs four parameters that are explained as follows.

Syntax: *asc2map --clone clone.map -a asciifile.txt result.map*

1. *--clone clone.map*: 'clone.map' is taken as clone map. This is a void map that contains the number of rows and columns, cell size, x and y coordinates and the data type. The number of rows and columns of the original ARC/INFO map given in the header must correspond with the number of rows and columns of the clone map.

2. *-a*: option used to convert ARC/INFO ascii files.

3. *asciifile.txt*: ascii file that comes from the ArcGIS process.

4. *result.map*: name of the resultant PCRaster map.

The input maps to run the model are classified in eight categories. The main files are described below and these categories are showed in Table 4.3.

Table 4.3 LISFLOOD input maps

GENERAL		
Map	**Name**	**Description**
Mask	area.map	Boolean map that defines model boundaries
TOPOGRAPHY		
Map	**Name**	**Description**
Local direction map	ldd.map	This file contains flow directions from each cell to its steepest neighbor (with value 1-9)
Grad	gradient.map	Slope gradient [m m^{-1}]
LAND USE		
Map	**Name**	**Description**
Land use	landuse.map	Map with land use classes
Forest	forest.map	Forest fraction for each cell. Values range from 0 (no forest at all) to 1 (pixel is 100% forest)
Direct runoff fraction	directrf.map	Fraction urban area for each cell. Values range from 0 (no urban area at all) to 1 (pixel is 100% urban)
SOIL		
Map	**Name**	**Description**
Texture 1	soiltex1.map	Soil texture class layer 1 (upper layer)
Texture 2	soiltex2.map	Soil texture class layer 2 (lower layer)
Soil depth	soildep.map	Depth to bedrock or groundwater [cm]
CHANNEL GEOMETRY		
Map	**Name**	**Description**
Channels	chan.map	Map with Boolean 1 for all channel pixels, and Boolean 0 for all other pixels on MaskMap
Channel gradient	changrad.map	Channel gradient [m m^{-1}]
Channel Manning	chainman.map	Manning's roughness coefficient for channels
Channel length	chanleng.map	Channel length [m]
Channel bottom width	channelbw.map	Channel bottom width [m]
METEOROLOGICAL VARIABLES		
Map	**Prefix**	**Description**
Precipitation	pr	Precipitation rate [mm day-1]
EO	e	Daily potential evaporation rate free water surface [mm
ESO	es	Daily potential evaporation rate, bare soil [mm day^{-1}]
ETO	et	Daily potential evapotranspiration rate reference crop [mm day^{-1}]
DEVELOPMENT OF VEGETATION OVER TIME		
Map	**Prefix**	**Description**
LAI	lai	Pixel-average Leaf Area Index [m2 m-2]
DEFINITION OF INPUT/OUTPUT TIMESERIES		
Map	**Name**	**Description**
Gauges	outlets.map	Nominal map with locations at which discharge timeseries are reported
Sites	sites.map	Nominal map with locations at which timeseries of intermediate state and rate variables are reported

The **mask map** ("area.map") defines the model domain (Van der Knijff et al., 2008). In order to avoid unexpected results, all maps related to topography, land use and soil were defined in such way that they do not contain missing values for each pixel that is "true" (has a Boolean 1 value) on the mask map. The same applies for all meteorological input and the Leaf Area Index maps.

Local drain direction map determine flow directions from each cell to its steepest down slope neighbour. This map was created by means of the 'pcrcalc' PCRaster application from the digital elevation model (DEM) with the following expression:

Syntax: *pcrcalc result.map = lddcreate(dem.map)*

Where *result.map* is the output of the application and *dem.map* is the digital elevation model.

Local drain directions are coded according to the following diagram (Van der Knijff et al., 2008):

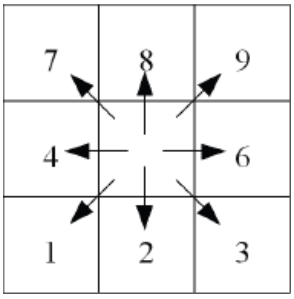


Figure 4.8 Local drain direction coding

Value '5' defines a cell without local drain direction (pit). The pit cell at the end of the path is the outlet point of a catchment. Therefore, value '5' of the local drain direction map coincides with the location of the discharge station 'H346' on the 'outlets.map'.

The **land use** feature class (ArcGIS format) obtained from MAG and CLIRSEN (see section 3) was converted to raster dataset and reclassified to numeric codes. These codes correspond to the crop coefficient and Manning's roughness codes presented on Table 4.4.

The soils map produced by DINAREN, is a feature class that contains on its attribute table the soils information. This information is based on the United States Department of Agriculture (USDA) classification. The field which has the 'ORDER' data was selected to rasterize this feature. Once more it was reclassified to get numeric codes that represent **soil texture**. These codes concord with the table codes of the soil texture files presented in Table 4.4.

All pixels that are "true" on the **channels map** have some valid (non-missing) value on each of the channel parameter maps. This is done since undefined pixels lead to unexpected behavior of the model.

The **meteorological forcing variables** and **Leaf Area Index** are defined in map stacks. A map stack is a series of maps, where each map represents the value of a variable at an individual time step (Van der Knijff et al., 2008). The name of each map is made up of a total of 11 characters: 8 characters, a dot and a 3-character suffix. Each map name starts with a prefix, and ends with the time step number. All character positions in between are filled with zeros '0' (see Figure 4.9).

Nonetheless, as meteorological variables are represented by a map per every time step, there are numerous ASCII files that should be converted to PCRaster maps. For that reason, it was necessary to create 'command files' that run on the Command prompt. They were firstly created in an excel sheet, copied to a text editor and then saved on the respective working directory (where the ascii files are) as 'run.cmd'. A section of the 'run.cmd' file used to create the precipitation map series is showed in Figure 4.9. By running the 'run.cmd' file, the respective maps for every time step were created.

In addition to the maps and time series, a number of model parameters are read by LISFLOOD through tables that are linked to the classes on the land use and soil (texture) maps. They are shown in Table 4.4.

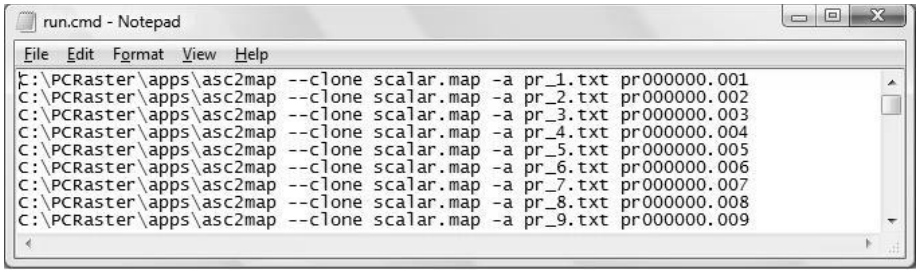


Figure 4.9 Conversion from ASCII precipitation files to precipitation PCRaster map series

Table 4.4 LISFLOOD input tables

LAND USE		
Table	**Name**	**Description**
Crop coefficient	cropcoef.txt	Crop coefficient for each land use class [-]
Manning's roughness	n.txt	Manning's roughness for each land use class [-]
SOIL TEXTURE		
Table	**Name**	**Description**
TabThetaSat1	thetas1.txt	Saturated volumetric soil moisture content layer 1 [-]
TabThetaSat2	thetas2.txt	Saturated volumetric soil moisture content layer 2 [-]
TabThetaRes1	thetar1.txt	Residual volumetric soil moisture content layer1 [-]
TabThetaRes2	thetar2.txt	Residual volumetric soil moisture content layer2 [-]
TabLambda1	lambda1.txt	Pore size index (λ) layer 1 [-]
TabLambda2	lambda2.txt	Pore size index (λ) layer 2 [-]
TabGenuAlpha1	alpha1.txt	Van Genuchten parameter α layer 1 [-]
TabGenuAlpha2	alpha2.txt	Van Genuchten parameter α layer 2 [-]
TabKSat1	ksat1.txt	Saturated conductivity layer 1 [cm day^{-1}]
TabKSat2	ksat2.txt	Saturated conductivity layer 2 [cm day^{-1}]

The **crop coefficient** values were taken from Table 2.5 of the Handbook of Hydrology (Maidment, 1992). This numbers are typical values expected under standard climatic conditions. The table gives mean crop coefficients for use with ETO for a number of different crops.

On the other hand, the soil water properties required by the soil texture tables were taken from Table 2. Hydrologic Soil Properties Classified by Soil Texture of the journal article Estimation of Soil Water Properties (Rawls et al., 1982).

4.5.2. Settings file

All file and parameter specifications have been defined in the settings file. The purpose of the settings file is to link variables and parameters in the model to in- and output files (maps, time series, tables) and numerical values (Van der Knijff et al., 2008). The settings file has an XML ('E**x**tensible **M**arkup **L**anguage') structure and it was edited using a dedicated XML editor.

The settings file is made up of the elements 'lfuser', 'lfoptions' and 'lfbinding'; each of which has a specific function. According to Van der Knijff (2008), the main function of each element is:

- lfuser: definition of paths to all files and main model parameters (calibration + time-related).
- lfbinding: definition of all individual files, and model parameters.
- lfoptions: switches to turn specific components of the model on or off.

Instead of writing the settings file completely, the template provided by LISFLOOD was edited. However, in order to use the template, the following requirements were met:

- All input maps and tables are named according to default file names
- All base maps, tables, meteorological inputs and Leaf Area Index maps are stored in a specific directory for each of these files
- An (empty) directory (for the output data files) exist

The present LISFLOOD version contains 24 process-related parameters (Van der Knijff et al., 2008). These are all defined in the 'lfuser' element, and default values are given for each of them.

4.5.3. Initialization of the model

The period of study is between January 1st. 2003 and December 31st. 2005 (3 years) but the model needs to have some estimate of the initial state of its internal state variables. The state of the soil at the beginning of that year was obviously unknown. However, it was solved by starting the simulation earlier. The meteorological and LAI maps of the year 2003 were copied twice and giving two years as a *warm-up* period, assuming that by the start of 2003 the influence of uncertain initial conditions is negligible.

Five internal state variables were set to zero at the start of the run: amount of water on the soil surface (*WaterDepthInitValue*), snow cover (*SnowCoverInitValue*), frost index (*FrostIndexInitValue*), interception storage (*CumIntInitValue*), and storage in the upper groundwater zone (*UZInitValue*). The initial value of the 'days since last rainfall event' (*DSLRInitValue*) was set to 1.

For the remaining state variables, LISFLOOD provides the possibility to initialize these variables internally, and these special initialization methods were activated by setting the initial values of each of these variables to a special 'bogus' value of *-9999*. Table 4.5 explains this special initialization.

Table 4.5 Special initialization methods

Variable	Description	Initialization method
ThetaInit1Value	initial soil moisture content upper soil layer (V/V)	set to soil moisture content at field
ThetaInit2Value	initial soil moisture content lower soil layer (V/V)	set to soil moisture content at field
LZInitValue	initial water in lower groundwater zone (mm)	set to steady-state storage
TotalCrossSectionAreaInitValue	initial cross-sectional area of water in channels	set to half of bankfull depth

4.5.4. Initialization of the lower groundwater zone

The time needed to initialize any storage component of the model is dependent on the average residence time of the water in it (Van der Knijff et al., 2008). Thus, the moisture content of the upper soil layer tends to respond almost instantly to meteorological forcing variables. As a result, relatively short warm-up periods are sufficient to initialize this storage component. However, the response of the lower groundwater zone is very slow and very long warm-up periods may be needed to avoid unrealistic trends in the simulations.

In order to avoid the need for excessive warm-up periods, LISFLOOD is capable of calculating a 'steady-state' storage amount for the lower groundwater zone. This *steady state* storage is very effective for reducing the lower zone's warm-up time.

The lower groundwater zone was initialized first doing a "pre-run" that was used to calculate the average inflow into the lower zone. This average inflow is reported as a map, which was then used in the actual run. Figure 4.10 shows a) the average inflow into the lower zone map for a wet period (January-June) and b) the average inflow into the lower zone map for a dry period (July-December).

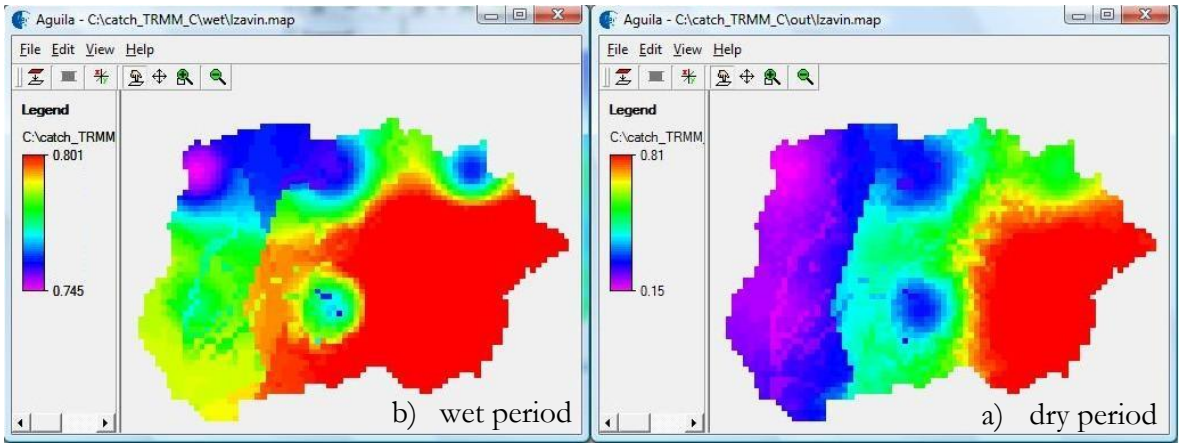


Figure 4.10 Average inflow into the lower zone map [mm/d] for wet and dry periods

4.5.5. Output generated by the model

Output is generated as either maps in PCRaster format or time series. Table 4.6 describes the output time series that are reported by default.

Table 4.6 Model default output time series

RATE VARIABLES AT GAUGES		
Description	**Units**	**File name**
channel discharge	m^3/s	dis.tss
NUMERICAL CHECKS		
Description	**Units**	**File name**
cumulative mass balance error	m^3	mbError.tss
cumulative mass balance error, expressed as mm water slice (average over catchment)	mm	mbErrorMm.tss
number of sub-steps needed for gravity-based soil moisture routine	-	steps.tss

4.6. Model Calibration

The procedure of adjusting the model input is necessary to match model output with measured field data for the selected period and situation entered to the model. This process is termed 'model calibration' (Rientjes, 2010). According to Van der Knijff et al. (2008), LISFLOOD model has five parameters that are used for calibration (see Table 4.7).

Table 4.7 Calibration parameters

Parameter	Description
UpperZoneTimeConstant	Time constant for water in upper zone [days]
LowerZoneTimeConstant	Time constant for water in lower zone [days]
GwPercValue	Maximum rate of percolation going from the Upper to the Lower Zone [mm/day]
b_Xinanjiang	Parameter in infiltration equation [-]
PowerPrefFlow	Parameter in preferential flow equation [-]

Calibration for the period 2003-2005 was done manually by 'Trial and Error' procedure. The observed values were taken from the H346 discharge station located at the outlet of the catchment (see section 3 Office collected data). For each run, parameter values in the settings file were changed. Firstly, parameters that control the base flow were fitted and after that, quick flow parameters were fitted. The results of the calibration were evaluated quantitatively by means of three objective functions described below.

1. The Root Mean Squared Error (RMSE)

RMSE is the measure of the differences between discharge values predicted by the model and the discharge values actually observed. This error is usually thought to be the best measure of error if errors are normally distributed (Rientjes, 2010).

The formula for RMSE reads:

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(Q_{measured} - Q_{calculated})_i^2\right]^{0.5}$$

Where:
$Q_{measured}$ is the measured discharge and
$Q_{calculated}$ is the modeled discharge

The maximum acceptable value of the calibration criterion was set during the calibration.

2. The Nash-Sutcliffe coefficient of efficiency (NS)

The Nash–Sutcliffe model efficiency coefficient is used to assess the predictive power of hydrological models and reads:

$$NS = 1 - \frac{\sum_{n=1}^{N}(Q_o^n - Q_m^n)^2}{\sum_{n=1}^{N}(Q_o^n - \bar{Q}_o)^2}$$

Where:

Q_o is observed discharge
Q_m is modeled discharge and
$\bar{Q}_o$ is the mean value of the observed discharge

NS values of 1 indicate perfect fits. NS values between 0.9 and 1 indicate that the model performs extremely well. Values between 0.8 and 0.9 indicate that the model performs very well while values between 0.6 and 0.8 indicate that the model performs reasonably well. Negative NS values indicate that the observed mean discharge is a better predictor than the model simulation.

3. The Relative Volume Error (RV_E)

The Relative Volume Error (RV_E) assesses the mass balance error between the observed and the simulated counterparts. The RV_E can vary between -∞ and +∞ but performs best when a value of '0' is generated since the accumulated difference between simulated and observed discharges is '0'.

The formula reads:

$$RV_E = \left(\frac{\sum_{i=1}^{n} Q_{sim(i)} - \sum_{i=1}^{n} Q_{obs(i)}}{\sum_{i=1}^{n} Q_{obs(i)}} \right) x\ 100$$

Where:

$Q_{sim(i)}$ is the simulated discharge
$Q_{obs(i)}$ is the observed discharge

A relative volume error between -5% and 5% indicates that a model performs well while a relative volume error between +5% and +10% and between -5% and -10% indicate a model with reasonable performance.

5. RESULTS AND DISCUSSION

5.1. Hydrograph simulation

In this study, the rain gauged observations are considered the true rainfall data. Consequently, the model was firstly run and calibrated with the rain gauged map timeseries. Figure 5.1 shows streamflow simulation results for the 3000 km^2 catchment in Babahoyo sub-basin. The simulation is compared to the hydrograph obtained from the observed discharges of the H346 hydrometric gauge.

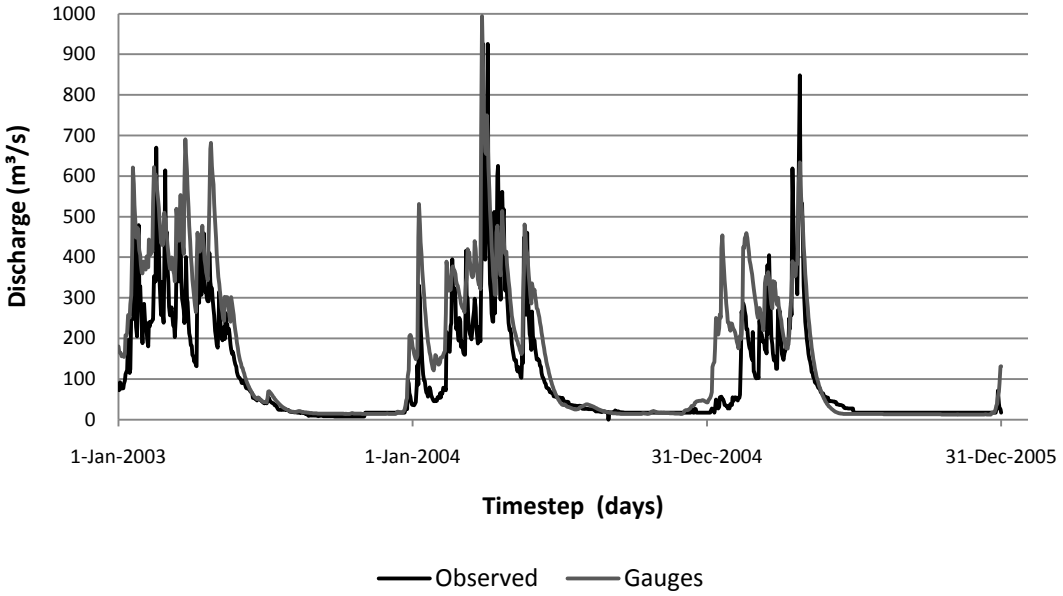


Figure 5.1 Hydrograph simulation (default parameters)

The simulated hydrograph is obtained with the default parameter values. The hydrograph present a poor match to the observed hydrograph, especially on the peaks of the quickflow. A large difference can be seen at the beginning of 2005. However, the baseflow of the two hydrographs nearly coincide.

Table 5.1 shows the introduced default parameters and the achieved objective functions. These values indicate the poor match between the simulated and the observed hydrographs.

Table 5.1 Default parameter and objective functions values

UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
10.00	1000.00	0.50	0.20	3.00	94.63	0.52	44.19

5.2. Sensitivity analysis

Different values for the calibration parameters were introduced to the model. This procedure was performed in order to quantify how sensitive the model is to changes in the calibration parameters. Moreover, sensitivity analysis was completed by changing one calibration parameter value at a time to avoid that the information content with respect to sensitivity of single values becomes obscured. The prior calibration parameter value ranges suggested by De Roo (2010) can be seen in Table 5.2. Moreover, for each parameter assessment, tables with the combination of parameter values and the resultant objective functions are shown. In those tables, the combination of parameters with the best objective function values and most similar hydrograph are in bold.

Table 5.2 Upper and lower bounds of calibration parameters

Parameter	Default value	Lower bound	Upper bound
UpperZoneTimeConstant	10	1	50
LowerZoneTimeConstant	1000	50	5000
GwPercValue	0.5	0	1.5
b_Xinanjiang	0.2	0.1	1
PowerPrefFlow	3	1	6

5.2.1. Effects on the UZTC parameter

Figure 5.2 shows that UpperZoneTimeConstant mainly controls the shape of the recession part of the hydrograph. Small values result in a fast, steep recession, whereas the slope of the falling limb is gentler for higher values. Table 5.3 shows that a UZTC value of '20' gives better objective functions.

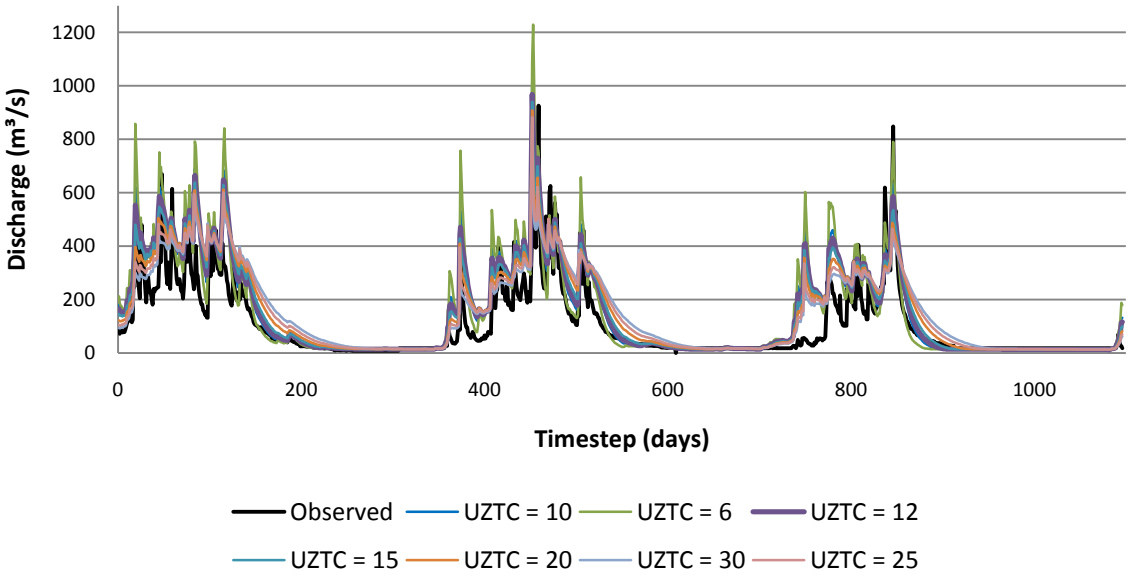


Figure 5.2 Sensitivity of the model to change in UZTC

Table 5.3 Parameters and objective functions for UZTC sensitivity analysis

RUN	PARAMETERS					OBJECTIVE FUNCTIONS		
	UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
default	10	1000	0.5	0.2	3	94.63	0.52	44.19
1	6	1000	0.5	0.2	3	106.34	0.39	44.06
2	12	1000	0.5	0.2	3	92.39	0.54	44.21
3	15	1000	0.5	0.2	3	90.64	0.56	44.20
4	**20**	**1000**	**0.5**	**0.2**	**3**	**89.75**	**0.57**	**44.13**
5	30	1000	0.5	0.2	3	91.30	0.55	43.92
6	25	1000	0.5	0.2	3	90.16	0.56	44.03

5.2.2. Effects on the LZTC parameter

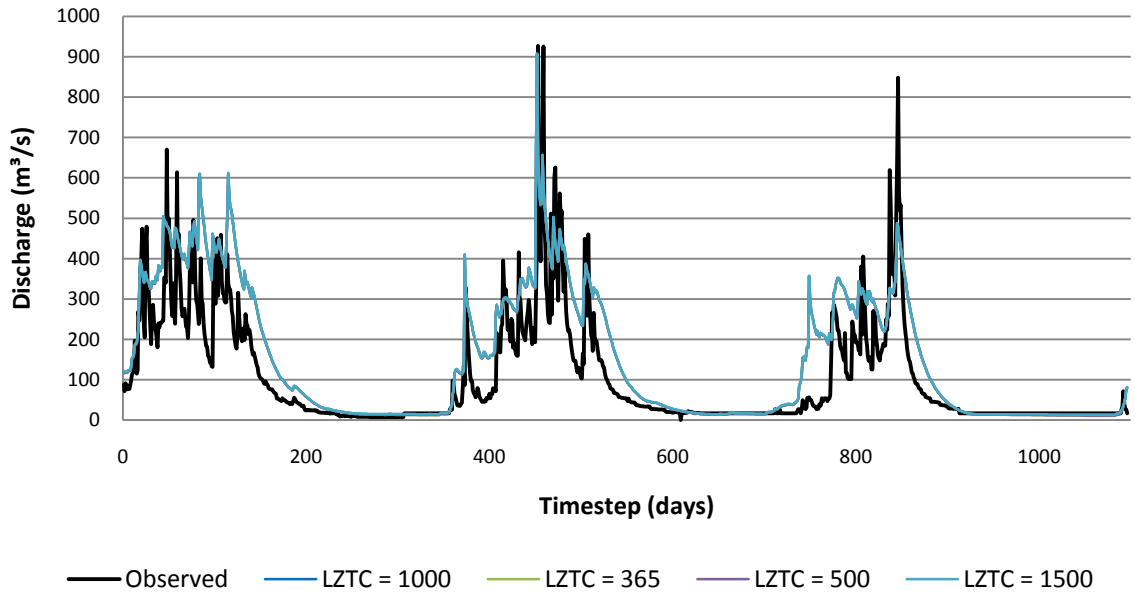


Figure 5.3 Sensitivity of the model to change in LZTC

LowerZoneTimeConstant controls much of the baseflow response. LZTC value of '365' gives better objective functions (see Table 5.4). Generally, this parameter did not affect largely the obtained hydrograph as it can be seen on Figure 5.3. Table 5.4 also indicates that the resultant objective functions are barely affected by the LZTC parameter value.

Table 5.4 Parameters and objective functions for LZTC sensitivity analysis

RUN	PARAMETERS					OBJECTIVE FUNCTIONS		
	UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
last	20	1000	0.5	0.2	3	89.75	0.57	44.13
1	**20**	**365**	**0.5**	**0.2**	**3**	**89.20**	**0.57**	**43.16**
2	20	500	0.5	0.2	3	89.32	0.57	43.38
3	20	1500	0.5	0.2	3	90.03	0.56	44.64

5.2.3. Effects on the GPV parameter

GwPercValue controls the baseflow behavior. Higher values results in large amounts of water in the baseflow section of the hydrograph. This effect is clearly illustrated on Figure 5.4.

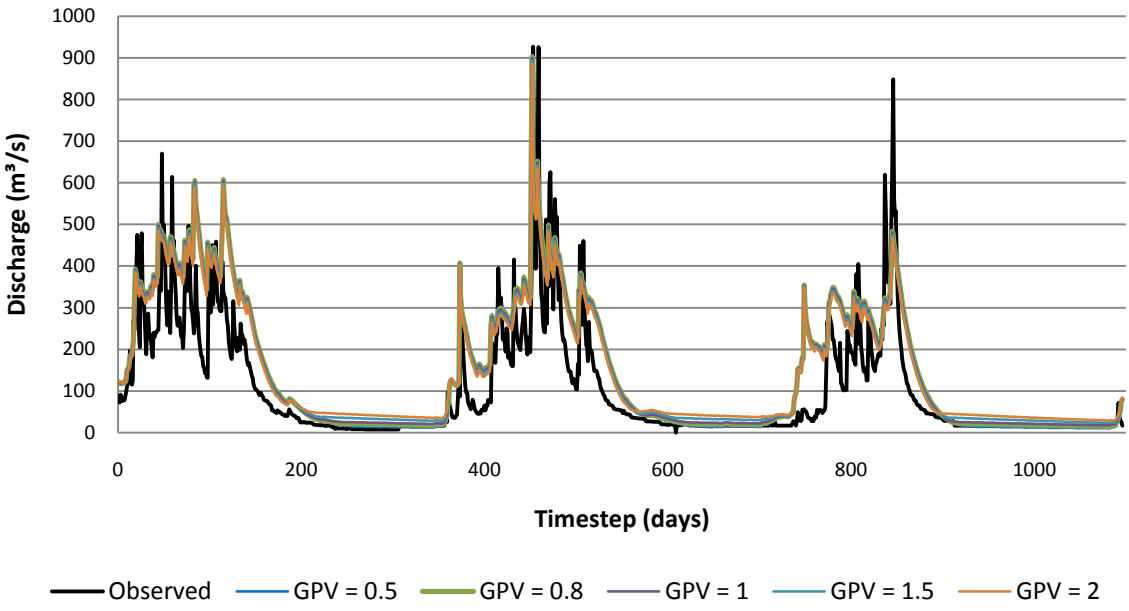


Figure 5.4 Sensitivity of the model to change in GPV

Table 5.5 indicates a small difference between the performance indicators obtained from runs 1 and 2. However, the resultant hydrograph with GPV value of '0.8' matches the observed baseflow better (see Figure 5.4).

Table 5.5 Parameters and objective functions for GPV sensitivity analysis

RUN	PARAMETERS					OBJECTIVE FUNCTIONS		
	UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
last	20	365	0.5	0.2	3	89.20	0.57	43.16
1	**20**	**365**	**0.8**	**0.2**	**3**	**87.75**	**0.59**	**43.56**
2	20	365	1	0.2	3	86.81	0.59	43.84
3	20	365	1.5	0.2	3	84.66	0.61	44.51
4	20	365	2	0.2	3	82.90	0.63	45.16

5.2.4. Effects on the bX parameter

The b_Xinanjiang parameter controls the infiltration. Increasing its value decreases the infiltration and thus rain rate becomes available for surface runoff. The parameter b_Xinanjiang affects surface runoff flow of the quickflow component. Differences of the obtained peaks are relatively small as Figure 5.5 suggests. The differences on the peaks are indicated in the circles of the figure when bX=0.2 and bX=0.01.

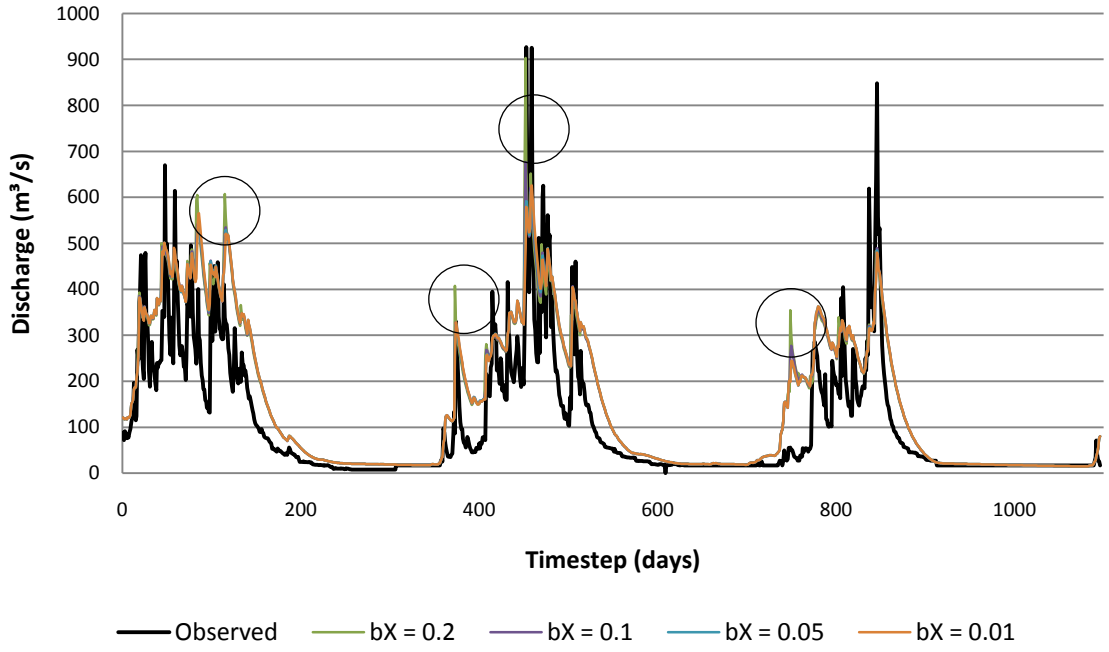


Figure 5.5 Sensitivity of the model to change in bX

A bX value of '0.05' gives the best objective functions as indicated in Table 5.6. According to Feyen et al (2007), this value is the lower bound for the b_Xinanjiang parameter. De Roo (2010) suggest that the lower bound for bX is 0.1 (see Table 5.2).

Table 5.6 Parameters and objective functions for bX sensitivity analysis

RUN	PARAMETERS					OBJECTIVE FUNCTIONS		
	UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
last	20	365	0.8	0.2	3	87.75	0.59	43.56
1	20	365	0.8	0.1	3	85.73	0.60	43.62
2	**20**	**365**	**0.8**	**0.05**	**3**	**85.46**	**0.61**	**43.58**
3	20	365	0.8	0.01	3	85.75	0.60	43.58

5.2.5. Effects on the PPF parameter

According to De Roo (2010), the lower and the upper bounds suggested for this parameter are 1 and 6 respectively (see Table 5.2) and the PowerPrefFlow parameter affects preferential flow of the quickflow component. The PPF parameter was changed from 0.5 to 4 and the resultant hydrographs indicate that the model is not very sensitive to this parameter for the catchment of study (see Figure 5.6). Nevertheless, PPF value of '3' gives better performance indicators as Table 5.7 shows.

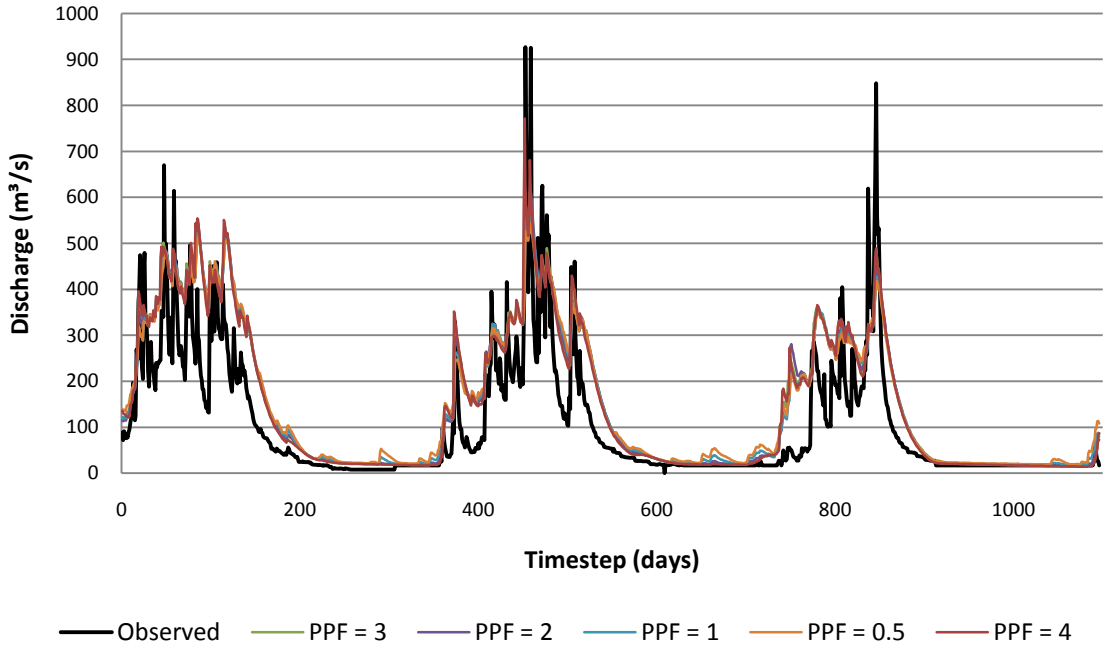


Figure 5.6 Sensitivity of the model to change in PPF

Table 5.7 Parameters and objective functions for PPF sensitivity analysis

RUN	PARAMETERS					OBJECTIVE FUNCTIONS		
	UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
last	**20**	**365**	**0.8**	**0.05**	**3**	**85.46**	**0.61**	**43.58**
1	20	365	0.8	0.05	2	86.73	0.60	43.81
2	20	365	0.8	0.05	1	87.04	0.59	44.62
3	20	365	0.8	0.05	0.5	87.00	0.59	46.41
4	20	365	0.8	0.05	4	86.16	0.60	43.53

5.2.6. Effects on the channel bottom width

LISFLOOD requires several maps that describe the channel geometry (see Table 4.3). In order to test the sensitivity of the model to the channel bottom width map, the model was run with different channel bottom width maps. In the procedure, all channel represented cells have a constant value and the upstream elements have the same channel width as compared to downstream channel elements. This setup only serves as a sensitivity analysis and the objective functions obtained from this testing are shown in Table 5.8.

Simulation results suggest that changing the values of the channel bottom width hardly shows an effect. Presumably effects would be observed where the channel geometry is especially significant such as simulations of flood events.

Table 5.8 Channel bottom width sensitivity

RUN	chanbw.map	RMSE(m^3/s)	NS	RVE (%)
1	5	85.852	0.604	46.359
2	10	85.920	0.603	46.360
3	25	86.061	0.602	46.361
4	50	86.223	0.600	46.362
5	1	85.774	0.604	46.359

5.3. Calibration results

For model calibration, rainfall timeseries from the rain gauged are used. Calibration by means of a Trial and Error procedure was performed and the assessment focused on qualitative and quantitative analysis. The measured and computed discharges were compared and adjustments were made to the five LISFLOOD calibration parameters to improve the match between the measured data and the computed output.

Figure 5.7 shows the result of the calibration process over the simulation period 2003-2005. Optimized parameter values are presented in Table 5.9 with the model performance indicators, Root Mean Square Error (RMSE), Nash-Sutcliffe coefficient of efficiency (NS), and Relative Volume Error (RVE).

Table 5.9 Optimized parameter values and objective function values after model calibration

UZTC	LZTC	GPV	bX	PPF	RMSE(m^3/s)	NS	RVE (%)
12	365	0.8	0.05	0.5	85.85	0.60	46.36

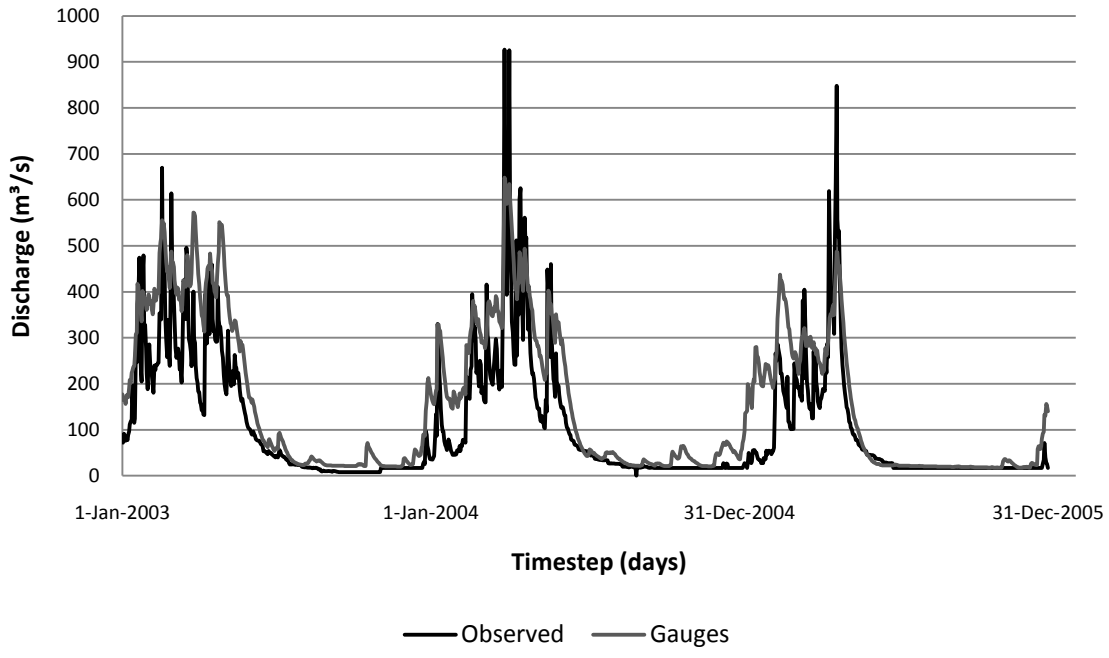


Figure 5.7 Hydrograph simulation (calibrated parameters)

Figure 5.7 shows a mismatch in the simulated baseflow in 2003 and 2004 while the simulated baseflow converge to the observed baseflow in 2005. Furthermore, several peaks and the falling limb of the calculated hydrograph are shifted forward. The large difference observed in the rising limb of the beginning of 2005 remains despite the calibration.

5.4. Comparison of TRMM 3B42 product and gauged data

LISFLOOD allows activating options to report additional output maps and time series. The 'repMeteoUpsGauges' option reports timeseries of meteorological inputs. These inputs are averaged by the model over the contributing area of the hydrometric gauging station. By using this option, the averaged 'precipitation' time series from gauged and TRMM input data were acquired. Table 5.10 presents various descriptive statistics of these two timeseries. This table shows rainfall underestimation by the TRMM data compared to the gauged data. The mean, maximum and sum of TRMM are lower than the respective values of the gauged data.

Table 5.10 Descriptive statistics for gauged and TRMM rainfall timeseries

Description	Gauged	TRMM
Mean	5.63	2.98
Standard Deviation	10.23	6.45
Minimum	0.00	0.00
Maximum	88.10	65.41
Sum	6171.31	3265.50
Count	1096	1096

Additionally, to compare the rainfall distribution between gauged and TRMM data, a daily rainfall histogram was obtained from the 'precipitation' timeseries of gauged and TRMM simulations. Figure 5.8 indicates large differences in observations frequency of rainfall depth between the two data sources, especially for the lower rainfall ranges.

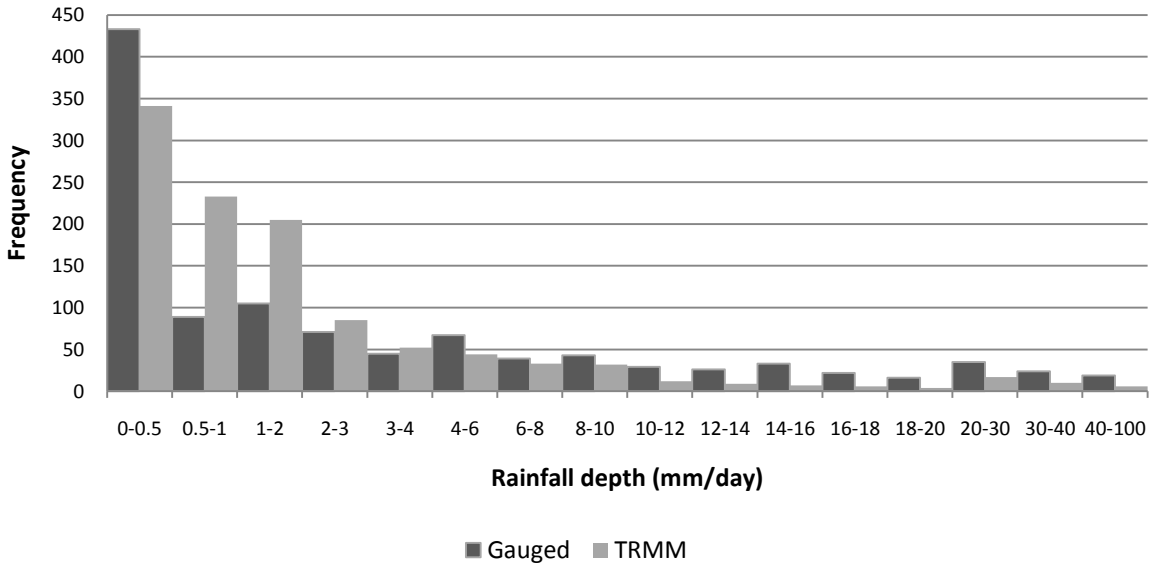


Figure 5.8 Daily rainfall distribution based on gauged and TRMM rainfall data (2003-2005)

Furthermore, the gauged and TRMM data were compared by means of double-mass analysis. Double-mass analysis tests the consistency of the TRMM records by comparing its accumulated mean rainfall with the concurrent accumulated mean rainfall from gauged stations. Figure 5.9 indicates changes in slopes that specify changes in the rainfall regime. In order to quantify these differences, a trend line was obtained and its equation suggests that the gauged data presents approximately 1.9 more rainfall than the TRMM product.

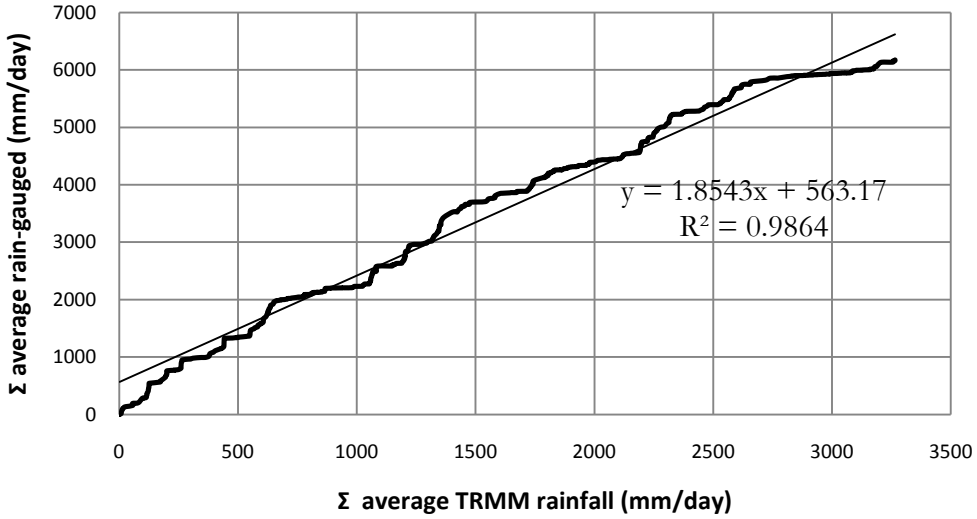


Figure 5.9 Double-mass curve between gauged and TRMM rainfall data

To define the volumetric difference between the two fields of rainfall (gauged and TRMM), a GIS procedure was applied. The 'CellStatistics' tool from ArcGIS was used with the following script: "CellStatistics_sa pr_1;pr_2;pr_3;… result SUM", where 'pr_i' are the raster rainfall maps, 'result' is the obtained raster and 'SUM' is the statistic type to be calculated. Subsequently, the sum of the pixels, which is the total amount of rainfall, was obtained for the gauged and TRMM datasets. Finally, the ratio between these two datasets was obtained. Total amount of water and the respective ratio are shown in Table 5.11.

Table 5.11 Total rainfall and ratio

	Total rainfall (mm)
Meteorological stations	17863219.66
TRMM 3B42 product	9450503.36
Ratio:	**1.89**

5.5. Rainfall representation and catchment responses

5.5.1. Rainfall representation

Rainfall variability is represented by a map stack in the model. This map stack is a series of maps where each map represents a rain field for the respective time step in mm/day. In this study, two rainfall representation approaches were applied. Rain gauged data was represented by the Thiessen polygons from five meteorological stations while TRMM 3B42 data was interpolated

using the inverse distance weighted (IDW) technique. Figure 5.10 shows these two rainfall representations for January 1st 2003.

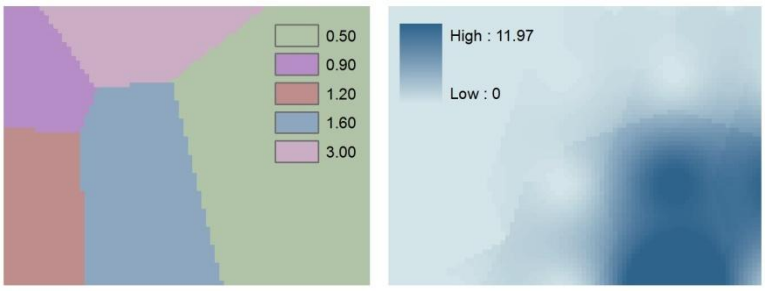


Figure 5.10 Thiessen polygons (rain gauged) vs. IDW (TRMM) [mm/day]

Rainfall represented by the meteorological stations has a single value in each Thiessen polygon of each map whereas rainfall represented by the processed TRMM product has diverse interpolated values on each pixel of each map (see Figure 5.10).

To assess the rainfall representation and the catchment responses from the model, the input water volume of the rain gauged and TRMM data should be similar. If this amount of input water is equal, the outputs of the model should be influenced only by the rainfall representation and not by the entered amount of water. To match the amount of water, the TRMM data set was corrected by multiplying the 1096 TRMM rainfall maps by the obtained ratio '1.890187' (see Table 5.11).

To evaluate the spatial difference between the summed rain-gauged map and the summed TRMM rain corrected map, the value of the summed gauged map was subtracted from the value of the summed TRMM corrected map on a cell-by-cell basis. Figure 5.11shows high differences values that range from -3250.83 to 7286.73 rainfall depth [mm]. These values indicate that numerous cells of the TRMM rainfall maps were greatly affected when the correction factor of 1.89 was applied.

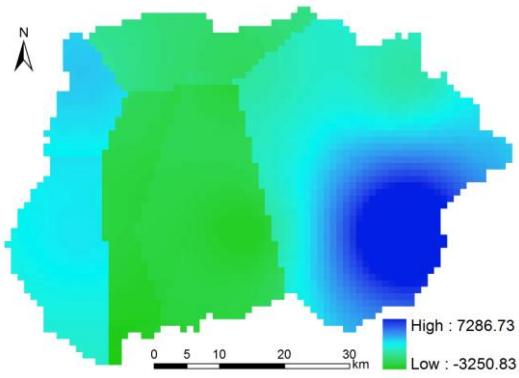


Figure 5.11 Spatial difference of summed rainfall maps [mm]

LISFLOOD model was separately run with the gauge rainfall maps and the corrected TRMM rainfall maps. Calibrated parameters obtained from the rain gauged input simulation are used in both. Before running the model, two additional options were activated in the settings files, 'repStateUpsGauges' and 'repRateUpsGauges'. These options report time series of model state

and rate variables respectively, averaged over the contributing area of the hydrometric gauge. Two state variables were analyzed, upper layer soil moisture and lower zone groundwater. In addition, two rate variables were analyzed, actual evaporation and actual transpiration. Finally, the simulated hydrographs are compared.

5.5.2. Catchment responses from gauged and TRMM rainfall inputs

Catchment responses from gauged and TRMM simulations are assessed to examine the influence of rainfall representation. Figure 5.13 presents the output averaged variables over the catchment and the input rainfall (P). Actual evaporation (AE) from the TRMM simulation presents higher values than the actual evaporation obtained from the gauges simulation during the dry seasons. The actual transpiration (AT) shows a roughly similar behavior for the two simulations but the AT from TRMM model is larger than the AT from the gauged simulation in the second half of 2005. Upper layer soil moisture (θ) illustrates a comparable performance for both models. However, the lower zone groundwater (LZ) from the TRMM model presents larger values than the lower zone groundwater values from the gauges model. Larger variations in the simulations are present especially in the dry period of 2005.

Figure 5.12 shows hydrograph simulation results when the corrected TRMM rainfall is used as model input. This hydrograph is compared to the simulation obtained when the gauges rainfall is the input. The rising and falling limb of the gauges simulation generally suggest more direct runoff than the TRMM simulation. However, the TRMM simulation presents larger baseflow (see Figure 5.12). These differences compensate and cause that Relative Volume Error (RVE) is relatively low (-6.61%).

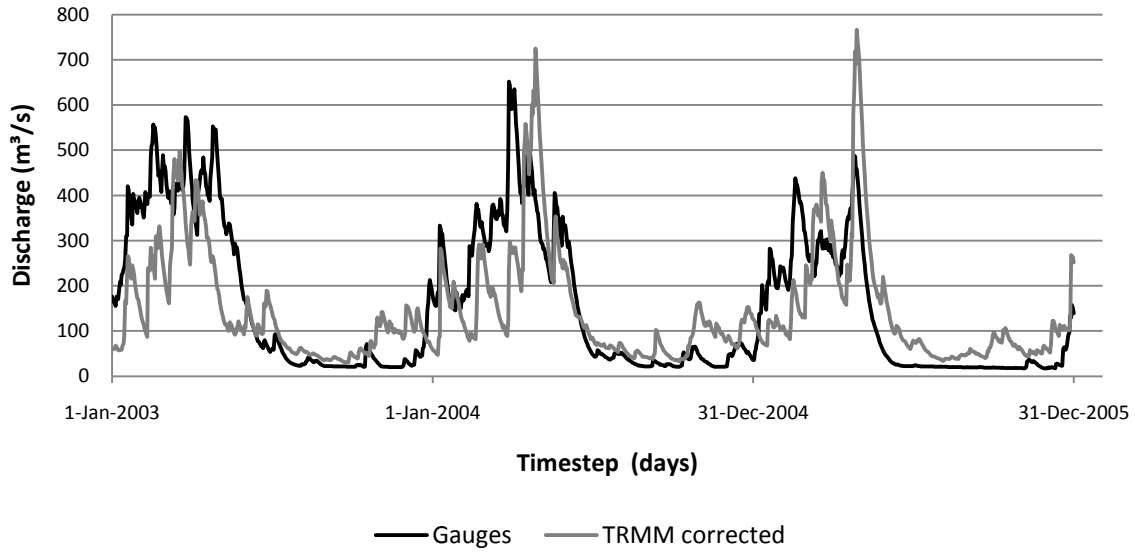


Figure 5.12 Hydrograph simulations (TRMM corrected)

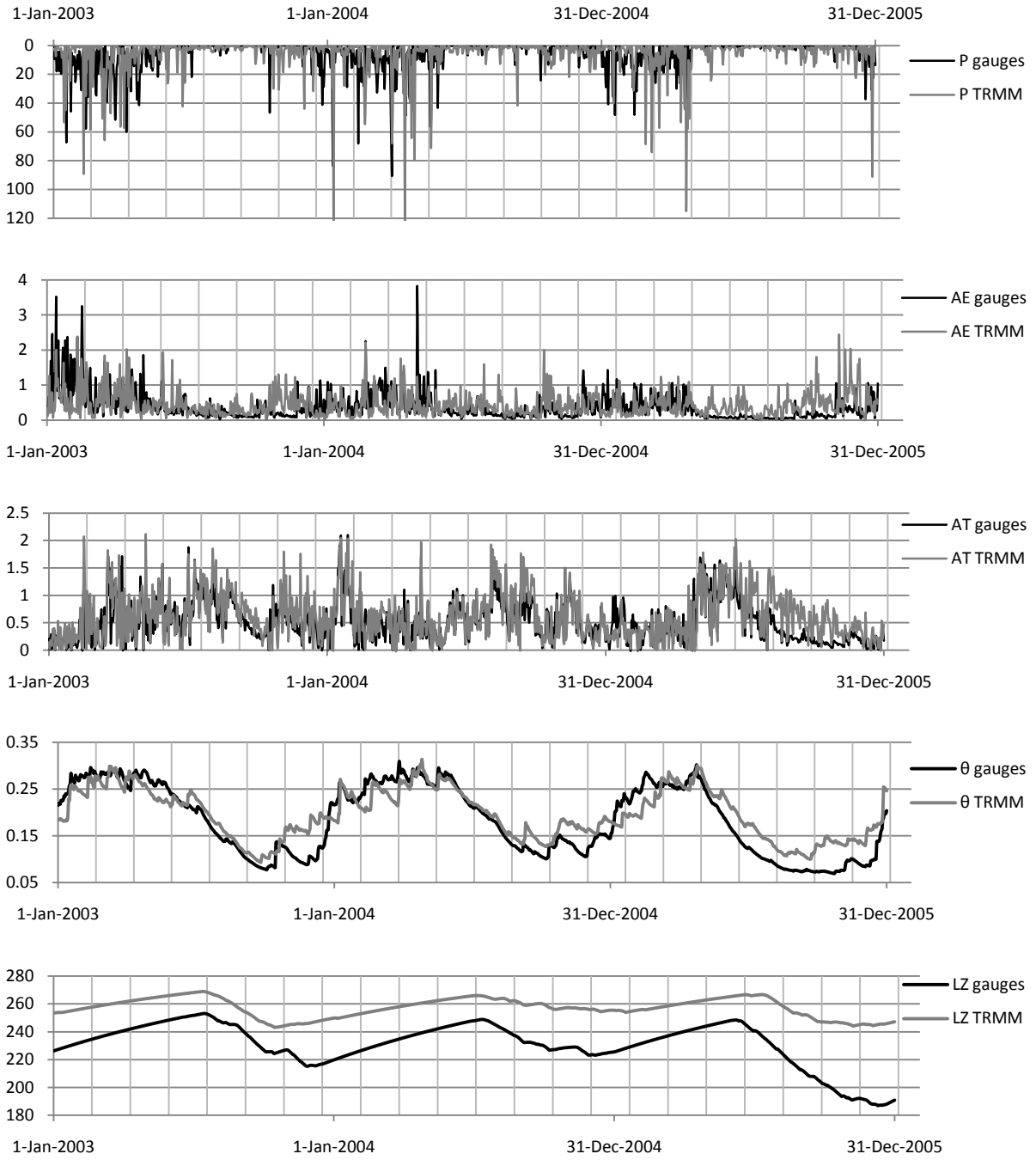


Figure 5.13 Catchment responses from gauged and TRMM rainfall inputs

Figure 5.13 shows fluctuations in the variables of the two models that generally coincide during the first two years and a half. Nonetheless, TRMM simulation presents higher amount of water than the gauged simulation for the rest of the period. Lower zone groundwater (LZ) is the unique variable where the TRMM simulation gives constantly larger values than the lower zone groundwater values of the gauges simulation.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

Three years (2003-2005) of data were used to perform a GIS and remote sensing based distributed model on a daily base to simulate runoff for a 3000 km^2 catchment in the Babahoyo sub-basin. This research serves as a pilot study for future applications in the Guayas River basin.

The LISFLOOD model was applied to simulate the runoff production from rainfall gauges. Runoff responses are in concordance with the rainfall input introduced to the model although the objective function values obtained during the model calibration are relatively low. The difference on the simulation results is probably because of the poor rainfall representation by low rain gauge network density.

The calibration of the LISFLOOD model is an important and vast task. It involves tuning parameters that cannot be obtained by direct measurements against observed discharges for the catchments. In this work a manual calibration was performed and results show that the Try and Error methodology was able to converge to the target hydrograph after numerous runs.

Rain rate observations retrieved from gauge and TRMM data are compared to assess the runoff of the study catchment. The rainfall comparison is assessed in terms of rainfall amount and spatial representation. Rainfall obtained from rain gauges in general produce much higher values compared to the TRMM 3B42 daily product.

TRMM rainfall represented by the Inverse Distance Weighted interpolation and gauged rainfall represented by the Thiessen polygons are compared. The assessment suggests that LISFLOOD model is sensitive to rainfall representation.

Remote sensing was used for acquiring rainfall and vegetation cover data. Furthermore, the applied GIS tools played a dominant role in obtaining raster data sets for numerous time series maps required by LISFLOOD. This study shows that integrate remote sensing together with the spatial data handling capabilities of GIS is significant to process data for distributed hydrologic modelling.

An unsuccessful attempt has been made to improve the evapotranspiration estimates by remote sensing. The lack of cloud free images, time and meteorological data constraints leaded to replace the remote sensing approach by the method used in this study.

6.2. Recommendations

This study provides relevant information about GIS based distributed rainfall-runoff modeling. Information about surface water resources availability on daily base on areas with high topographic variation such as the Babahoyo sub-basin is required to study for instance floods processes and for runoff simulation and forecasting. As such, similar studies can be undertaken for each sub-basin of the Guayas River basin to have more detailed information on the rainfall-runoff relation over the basin.

Manual calibration by optimizing parameters in this study proved to be very time consuming. Feyen et al. (2007) has demonstrated the capability of Shuffled Complex Evolution Metropolis (SCEM-UA), an automatic Bayesian parameter inference algorithm based on Markov chain Monte Carlo methods, to automatically calibrate the LISFLOOD model against daily discharge observations. Therefore, it is recommendable to apply SCEM-UA for automatic calibration of LISFLOOD model.

Calibration and validation against additional observations of the catchment, such as soil moisture, actual evapotranspiration or groundwater is an option to further test the model performance. Bergstrom et al. (2002) suggests that multi-variable parameter estimation can increase confidence in hydrological modeling.

A different specific method to study the catchment runoff response would be to divide the model calibration into two separated periods, a model which better represents the wet season (from December to June), and another that represents the dry season (from June to December).

In the Guayas River basin, there are 69 meteorological stations distributed over the area. The information obtained from these rain gauges can serve to compare the TRMM 3B42 rainfall product estimates on an area of 32000 km^2. By this procedure, rainfall amount and spatial representation could be assessed with extra field data over a larger surface.

To retrieve evapotranspiration estimates, advanced remote sensing techniques that involve the use of Normalized Difference Vegetation Index (NDVI) and large scale meteorological models can be applied.

Module 2 of the Guayas River Basin Project is studying the soils of 16 Cantons located in the basin. Among other initiatives, the module serves for obtaining the hydraulic properties of soils through fieldwork campaigns and laboratory processes. This information could help to improve the model setup and thus, describe on a better approach the soil moisture behaviour that takes place into the two soil compartments of LISFLOOD model.

The basin with the most significant flood damage in Ecuador is the Rio Guayas basin (Buckalew et al., 1998). Although this model's primary output product is channel discharge, the output of this rainfall-runoff model can be used to provide the initial conditions of a flood simulation in the Guayas River basin.

LISFLOOD model has been tested only on European catchments. This work presents the possibility of the model to be applied in other catchments located in The Andes region. This gives additional options to assess the surface water resources in the region on a distributed fashion.

LIST OF REFERENCES

Abbott, M. B. and J. C. Refsgaard (1996). Distributed Hydrological Modelling. Dordrecht, Kluwer Academic Publishers.

Allen, R. G., W. O. Pruitt, et al. (2005). "Estimating Evaporation from Bare Soil and the Crop Coefficient for the Initial Period Using Common Soils Information." Journal of Irrigation and Drainage Engineering **131**(1): 14-23.

Almeida, A., W. Armas, et al. (2009). Modulo 5 Uso de la Tierra. Proyecto Cuenca del Río Guayas. Quito, CLIRSEN**:** 38.

Bergstrom, S., G. Lindstrom, et al. (2002). "Multi-variable parameter estimation to increase confidence in hydrological modelling." HYDROLOGICAL PROCESSES **16**: 413–421.

Beven, K. J. (2000). Rainfall-Runoff Modelling The Primer. West Sussex, John Wiley & Sons Ltda.

Buckalew, J., M. James, et al. (1998). Water Resources Assessment of Ecuador, US Army Corps of Engineers, Mobile District and Topographic Engineering Center**:** 83.

Buckalew J., J. M., Scott L., Reed P. (1998). Water Resources Assessment of Ecuador, US Army Corps of Engineers, Mobile District and Topographic Engineering Center**:** 83.

CLIRSEN Magazine (2009). Teledetección 2009. Quito. 68.

De Roo, A. (2010). LISFLOOD Distributed Water Balance and Flood Simulation Model. Setting up the model.

De Roo, A., C. Wesseling, et al. (2000). "Physically based river basin modelling within a GIS: the LISFLOOD model. Hydrological Processes." 14.

Encalada, M., H. Castanier, et al. (1997). Proyecto de Fortalecimiento y Ampliacion de los Servicios Basicos de Salud en el Ecuador (FASBASE). Quito.

FAO (1990). "Irrigation and Drainage Paper No. 56 Crop Evapotranspiration."

Feyen, L., J. A. Vrugt, et al. (2007). "Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model." Journal of Hydrology **332**(3-4): 276-289.

Gonzales, A. (2010). Estimation of actual evapotranspiration and crop coefficient in the REMENDHUS area, Spain. International Institute for Geo-information Science and Earth Observation ITC. Enschede. **Master of Science in Geo-information Science and Earth Observation:** 61.

Jain, M. K., U. C. Kothyari, et al. (2004). "A GIS based distributed rainfall-runoff model." Journal of Hydrology **299**(1-2): 107-135.

Linsley, R., M. Kohler, et al. (1982). Hydrology for Engineers. United States of America, McGraw-Hill.

Maidment, D. R. (1992). Handbook of Hydrology. United States, McGraw-Hill.

NASA (2010, November 03). "Daily TRMM and Others Rainfall Estimate (3B42 V6 derived)." from http://mirador.gsfc.nasa.gov/collections/TRMM_3B42_daily__006.

NASA (2011). "Aqua Project Science." from http://aqua.nasa.gov/about/.

NASA (2011). "TERRA (EOS AM-1) - About Terra." from http://terra.nasa.gov/About/.

Peñaherrera, E. (2009). Modulo 3 Clima e Hidrología. Proyecto Cuenca del Río Guayas. Quito, CLIRSEN**:** 26.

Rawls, W., D. Brakensiek, et al. (1982). "Estimation of Soil Water Properties." American Society of Agricultural Engineers **25**.

Rientjes, T. (2010). Modelling in Hydrology. Enschede, University of Twente
Fac. of Geo-information Science and Earth Observation (ITC)
Department of Water Resources**:** 240.

US Department of the Interior and USGS (2010, May 20). "MCD15A2." from https://lpdaac.usgs.gov/lpdaac/products/modis_products_table/leaf_area_index_fraction_of_photosynthetically_active_radiation/8_day_l4_global_1km/mcd15a2.

USGS. "MODIS Land Data Products and Services." from https://lpdaac.usgs.gov/lpdaac/products/modis_overview.

Van der Knijff, J. (2008). LISVAP Evaporation Pre-processor for the LISFLOOD Water Balance and Flood Simulation Model. Revised User Manual. Luxembourg, Joint Research Centre, European Commision. **ISSN 1018-5593:** 31.

Van der Knijff, J. and A. De Roo (2008). LISFLOOD Distributed Water Balance and Flood Simulation Model. Revised User Manual. Luxembourg, Joint Research Centre, European Commision. **ISSN 1018-5593:** 109.

Van Deursen, W. and C. Wesseling (1991). PCRaster version 2 Manual. Utrecht.

Wagner, S., H. Kunstmann, et al. (2009). "Water balance estimation of a poorly gauged catchment in West Africa using dynamically downscaled meteorological fields and remote sensing information." Physics and Chemistry of the Earth, Parts A/B/C **34**(4-5): 225-235.