Modelling a catchment and study site in Andhra Pradesh by applying a rainfall – runoff model developed for West Bengal

Bachelor Thesis performed April – June 2013

R.J.A. (Rinse) Wilmink

1 August 2013

Modelling a catchment and study site in Andhra Pradesh by applying a rainfall – runoff model developed for West Bengal

Bachelor Thesis - Final report

This report contains the bachelor thesis of Rinse Wilmink, the last course of the Bachelor Civil Engineering at the University of Twente. To compose this report, the structure of Gillet, Hammond, & Martala (2009) is used. The research has been performed at the Australian National University, Canberra - Australia, April – June 2013.

Author:	R.J.A. (Rinse) Wilmink
Student number:	s1095153
Contact details:	r.j.a.wilmink@student.utwente.nl
	+31 6522 72 72 4
Date:	Thursday, 1 August 2013

University of Twente, Enschede – The Netherlands

Supervisor:	J. (Joep) Van der Zanden, MSc
Contact details:	j.vanderzanden@utwente.nl
	+31 53 489 4038
Faculty:	Engineering Technology
Department:	Water Engineering & Management

Australian National University, Canberra – Australia

Supervisor:	Dr. B.F.W. (Barry) Croke
Contact details:	barry.croke@anu.edu.au
	+61 2 6125 0666
Department:	Fenner School of Environment and Society
	Integrated Catchment Assessment and Management Centre, iCAM

Summary

In Andhra Pradesh (state of India) watershed development programs (WSD) have been undertaken to improve the livelihood of farmers. Most farmers in India are poorly educated, have bad healthcare and only have small landholdings. The cropping for these farmers is therefore very difficult, which is reinforced by the shift in timing of the monsoon, the increased dry spells and the population growth.

Andhra Pradesh has a semi-arid climate, with an average yearly rainfall of 450 – 600 mm, falling mostly between July and October. Hydrologically, the catchment area of Andhra Pradesh consists of uplands (''Tarh'', not suitable for cropping), medium uplands (only suitable for cropping when terraced and bunded, then called ''Baidh''), and lowlands (suitable for cropping). The population growth is forcing farmers to crop outside the lowlands, shifting to the medium uplands, which are much dryer.

The terms Tarh and Baidh are used in a previous WSD study in West Bengal. Due to the different languages used in India, these terms are likely to be different from the terms used in Andhra Pradesh. However, for consistency, these terms will be used for the catchment modelling in Andhra Pradesh too.

This bachelor thesis is a small part in a WSD study into the effects of WSD policies in relation to the investments level of return on the meso-scale. The purpose of this research is to:

''Apply the model developed for the West Bengal study site to the study sites for a second project in Andhra Pradesh (India), extending the application to a larger scale.''

The research started with the original model developed for West Bengal by Groeneveld (2012). After a data analysis, which gave information about the input data, the individual rainfall gauges are chosen as the rainfall input data to continue this research. The original model has been modified based on performance indicators and a visual interpretation of a gauged catchment area ($\sim$2750 km^2) to create a model that generates a better representation of the catchment's hydrology. This modification included a change in model structure as well as in the model processes. Finally this model has been applied to an ungauged study site ($\sim$200 km^2).

The modified structure and processes represented in the model, based on visual interpretations of the catchment and study site using Google Earth, and applying a simple approach of new model processes, gave a better representation of the catchment's hydrology than the original model. The Nash-Sutcliffe coefficient and Relative Volume Error of the generated runoff by the model over the period of study are 0.66 and -15.8 % respectively. The underestimation by the model was expected and could be justified. Because of a lack of good quality, high resolution and high frequency data, only indications about possible problems with the data could be made where a mismatch of the modelled and measured runoff occurred. Also, scarce information of the catchment's area characteristics led to many assumptions, mostly derived from related studies. However, several possible data problems could be identified. These identified issues were able to clarify some observed errors in the data. This all resulted in recommendations and additional research themes about the data, model processes and area characteristics.

Applying the modified calibrated version of the model on the ungauged Gooty study site required a visual inspection of the study site, which led to a small change in model structure. The rainfall-runoff

coefficients of the gauged and ungauged areas gave comparable results. The difference in modelled rainfall-runoff coefficients could mostly be related to the change of input data between the gauged catchment and the ungauged study site.

Finally, the model developed in this research has approved capability of modelling the catchment and study site in Andhra Pradesh. These model codes (the code for the catchment and the slightly different code for the study site) can be used for estimating, and possibly measuring, the effects watershed developments in this region of India, trying to improve the livelihood of farmers.

Preface

After three years studying, the bachelor Civil Engineering at the University of Twente is almost completed. This research project, the bachelor thesis, is the last course of the bachelor. It was a great pleasure to perform this research project abroad, in Canberra - Australia. Therefore I want to thank Dr. Martijn Booij of the Water Engineering and Management department for usage of his network in water management and hydrological modelling.

I also want to thank Dr. Barry Croke, of the Australian National University – Integrated Catchment Assessment and Management Centre (iCAM) and the other colleagues at iCAM; they were a group of great people, which was a pleasure to work with. Dr. Croke provided the necessary information to compose this research, and was the supervisor of the thesis in Australia. He acted as a council for new ideas and taught me other views on the usage of data, especially when little data is available.

Supervisor at the University of Twente is Joep van der Zanden MSc. Joep guided me through the thesis and examined this final report. He also gave critical, detailed and clear feedback on concept sections of this report during the research which helped me to compose this final report, wherefore I would like to thank him.

Rinse Wilmink,

Enschede, 1 August 2013

Table of contents

1. Introduction

1.1. Background

The demand for food worldwide is increasing every year due to increasing world population, increasing welfare or increasing crop failures. Most farmers in India are poorly educated, have inadequate healthcare and only have small land holdings. Especially because of the lack of education and substantial land holdings, cropping is very difficult. This problem is reinforced by the shift in timing of the monsoon, the increased dry spells in the monsoon season and the population growth which forces farmers to crop outside the low laying wetlands called ''lowlands''; shifting to the much dryer medium uplands which are less suitable for rice cropping, the major crop.

In order to overcome the problems with these basic necessities, the government in India supports watershed management research in order to improve the livelihood of farmers, increase their knowledge about cropping and, at the end, increase the food production for the Indian population.

Figure 1 - Location Andhra Pradesh (Google Maps)

Watershed Development programs have been introduced to ensure the sustainability of the surface and groundwater resources, and to improve the livelihoods of farmers. The Australian Centre for International Agricultural Research (ACIAR) funds research to Watershed Development (WSD) programs in rain fed dry land agriculture in India (Syme, 2009). In West Bengal (state of India) a WSD study has already been carried out at the micro-catchment level (up to 500 hectares) in the Pogro catchment.

1.2. Watershed development program in Andhra Pradesh

For the current project: ''Impacts of meso-scale Watershed Development in Andhra Pradesh (India) and their implications for designing and implementing improved WSD policies and programs'', the focus is on the larger meso-basin level scale. Here, the aim is studying the effects of WSD policies in relation to the investments level of return (Syme, 2009).

Andhra Pradesh is one of the 28 states of India. It is located on the south-eastern coast of India, see Figure 1. Its total area is 275,045 km^2 with a population around 85 million inhabitants (Government of Andhra Pradesh, 2011). Andhra Pradesh has a semi-arid climate, with an average yearly rainfall of 450 – 600 mm, mostly falling between July and October (World Weather Online, 2013). The landscape of Andhra Pradesh is a hilled area that consist of drainage lines, and land near streams comprising lowlands, medium uplands and uplands (''Tarh'') (Cornish, Kumar, & Khan, 2010). From a historic point of view, the only area that is suitable for cropping during the monsoon season (without making any changes to the area), are the lowlands located in the wet part of the area. The medium

uplands and uplands are too dry and too steep for cropping; these land classes cannot hold the water in natural conditions and are therefore not suitable for irrigation.

Hydrologically, the uplands are recharge areas without cropping. The lowlands are the discharge areas for seasonally recharged groundwater. The medium uplands lie between the uplands and the lowlands, and are a discharge area in only very wet years. Because of the population growth, farmers are forced to shift to the medium uplands, which are terraced and bunded to make them suitable for cropping. This area of medium uplands made suitable for cropping is called ''Baidh'' (Cornish, Kumar, & Khan, 2010). The terms Tarh and Baidh are derived from West Bengali, the language spoken in West Bengal. Due to the different languages used in India, these terms are likely to be different from the terms in Andhra Pradesh. However, for consistency, the same terms (Tarh and Baidh) will be used for modelling the catchment areas in Andhra Pradesh.

The work for the WSD study in Andhra Pradesh is focused on a 150 to 250 km^2 scale - called meso-scale in the project title, and corresponding roughly to the scale of a Mandel (smallest scale of the 5 levels of the government in India - National, State, District, Block and Mandel). The meso-scale project has a greater social focus, looking at the impacts of watershed development both locally, as well as on downstream users. As such, the project has an interdisciplinary team, comprising of social scientists, economists and biophysical scientists of different organisations.

The main goal of this WSD study is to build an integrated model for each of the two study sites in Andhra Pradesh, using a Bayesian network approach, with information gathered from a number of surveys the villagers have taken part in. The goal for the hydrology group within the project is to provide the input necessary to capture the hydrologic processes of the study sites, and use it to study the impact of changes in climate and the effects of watershed development. This involves modelling the study sites and validating the modelled output with observations.

1.3. Integrated Catchment Assessment and Management Centre (iCAM)

This bachelor thesis was carried out at the Integrated Catchment Assessment and Management Centre, iCAM. iCAM is a department of the Australian National University, Canberra Australia. Their mission is "to develop and integrate the knowledge required to clarify management and policy options for sustaining vital water and related resources. It is underpinned by targeted research in hydrology, ecology and socio-economies, and by interdisciplinary projects co-designed with interest groups.'' (Jakeman, 2013).

iCAM does research to decision support; integrated assessment; hydrological and ecological modelling and sensitivity and uncertainty assessment. Most projects are in the scope of modelling and decision support. iCAM has developed several decision support systems and watershed models as results of their scientific research.

iCAM sometimes participates in research projects funded by the ACIAR along with other funding groups. The ACIAR (Australian Centre for International Agricultural Research) is an Australian governmental organisation working on the government's development cooperation programs and is trying to improve the agriculture in Australia as well as in developing countries (ACIAR, 2003). This WSD study is co-funded by the Australian and Indian governments, with the purpose of improving livelihoods in poor areas of the world.

2. Research goal and methodology

2.1. Purpose bachelor thesis

This bachelor thesis is a small part in a WSD study that aims to improve the livelihoods of farmers in Andhra Pradesh, India (see Introduction). The aim of the BSc project is to adapt the model developed for the two study sites in the Pogro catchment (West Bengal) to a study site of a catchment in Andhra Pradesh. Because every study site has unique dimensions, natural conditions and climate, up-scaling, modification and calibration of the model is necessary.

The model is available in FORTRAN, R and Matlab. The model have to be up scaled (because of the meso-scale in this project studying WSD instead of the micro-scale at the previous WSD project) and tested on an Andhra Pradesh study site to validate whether changes in the model structure are needed because of the model performance. Ultimately, the BSc project has to provide inputs to the integrated Bayesian Network model. Summarized, the purpose of the bachelor thesis is:

''Applying the model developed for the West Bengal study site to the study sites for a second project in Andhra Pradesh (India), extending the application to a larger scale.''

2.2. Research questions

To achieve this purpose, several research questions are formulated. Each research question matches a phase in this BSc project.

1. How does the original model perform on a ~2750 km^2 gauged catchment in Andhra Pradesh at a daily time step?
2. Which modifications can be made to the original model to increase the model performance on a gauged catchment (~2700 km^2) in Andhra Pradesh?
3. How does the modified model (developed in phase 2) perform on an ungauged study site in Andhra Pradesh (~200 km^2) and which modifications are advised to improve the model performance?

2.3. Report outline

This research has been performed in several phases. The sequence and methodology of these phases is described in section 2.4, and also forms the outline of this report. After the explanation of the phases and methodology, the process of research including the results are shown per phase. This report outline per phase is explicitly chosen because results achieved in earlier phases are the foundation blocks of the forthcoming phases in this research. After the research process including the results, a discussion and conclusion about the results has been done, ending with a future work section.

2.4. Methodology

The method used for performing this research is extracted from Bennet et al. (2012). This method can be used in several environmental modelling purposes and is suitable for this hydrological modelling purpose.The method consist of five steps:

1. Have a clear idea of the modelling purpose.
2. Check the data available
3. Visual performance analysis

4. Selecting performance criteria
5. Refinements of the model

Step one has already been performed in section 1 and 2. The second step will be described in section 4; step 4 is carried out in section 5. Step 3 and 5 are performed in sections 4, 5 and 6 because refinements to the model may entail revision of the model structure and processes, so the procedure may require additional cycles (Bennet, et al., 2012).

2.4.1 Phase 0; Understanding of the model

The first phase consists mainly of understanding the model. The BSc project starts with the model developed by Groeneveld (2012), named as the original model. This model consists of several scenarios and calibrated parameters, further explained in section 3. These elements first have to be removed from the model to use a top-down modelling approach (start with a simple model based on observations, as described in e.g. Littlewood, Croke, Jakeman, & Sivapalan, 2003).

2.4.2 Phase 1; Applying the cleaned model on gauged catchment Andhra Pradesh

The study site in Andhra Pradesh does not have gauged streamflow data available. Therefore, the model is first being tested on a gauged catchment adjacent to the study site with the available data. Before the model run on Andhra Pradesh (section 4.2), a data analysis will be performed (step 2 of the method, section 4.1) to examine the general behaviour of the data.

The catchment (and the study site adjacent the catchment) in Andhra Pradesh has a different climate than the Pogro catchment. In the Andhra Pradesh catchment there is less rainfall. Also the time step of the data differs, therefore the model needs to be adjusted to the new catchment area. The model scale has to be changed from 4 km^2 to approximately 2750 km^2 for the gauged catchment and to ~200 km^2 for the study site. At the end of this phase, the model is adjusted to the catchment of Andhra Pradesh and ready for model structure changes or process modifications.

2.4.3 Phase 2; Modification of the original model

This step is the most important phase in this research. Main activities during this phase are to think about what modifications need to be made to the model to fit the catchment area in Andhra Pradesh better, implement these modifications, and check whether the model performance has improved by comparing the modelled runoff with the measured runoff. At the end of this phase, a calibration of the model parameters on this catchment is done in order to start with a model that fits the processes and structure of the catchment area, and probably the study site, as good as possible so that impacts of watershed development can be investigated.

2.4.4 Phase 3; Applying the model on an ungauged study site in Andhra Pradesh

After testing and calibrating the model on the gauged catchment in Andhra Pradesh, this model is applied on the ungauged study site of the WSD study. Based on the differences in land use between the gauged catchment and ungauged study site, several extra modifications will be explored and tested. To check the model performance of the extra modifications, a comparison is made with the rainfall – runoff coefficient of the gauged catchment of Andhra Pradesh.

Because of the lack of data about geographical area characteristics, land class characteristics and bad quality input data; many assumptions are made to create a hydrological model that represents the area as closely as possible.

3. Original model by Groeneveld

The research started with a model that has previously been calibrated for the Pogro study site (~ 4 km^2) by Groeneveld (2012). This model, named as the original model, is a modification of the IHACRES (Identification of unit Hydrographs And Component flows from Rainfall, Evapotranspiration and Streamflow data) rainfall runoff model of Jakeman, Littlewood, & Whitehead (1990) including a catchment moisture deficit (CMD) module (Croke & Jakeman, 2004, 2005). This model is able to adequately generate the runoff of the Pogro catchment. However, the model could not reproduce the hydrological processes within the catchment because of the existence of a surface layer. The latter is important to measure effects of WSD within a catchment. Most WSD effects are based on the change of surface storages to improve water retention.

The existence of a surface layer is important in WSD because WSD studies are focussed on improving water recources, and adjusting the surface storage through construction of pits, ponds and bunds. It is therefore necessary to include the hydrological processes operating in the surface layer within the model. In the Pogro model using the CMD module, the surface layer acted as a buffer between the CMD module and the rainfall data (the CMD module needs direct input from rainfall data), causing a deactivation of the infiltration in lower landclasses. As a result, the model was unable to capture the subsurface streamflow through the shallow aquifer properly (Groeneveld, 2012).

Groeneveld (2012) replaced the CMD module with a shallow aquifer that interacts with the surface module and produces a subsurface flow with the behaviour of a single reservoir unit hydrograph. The subsurface flow produces exfiltration to the surface storage which will be distributed to the lowlands first, proceeding uphill.

Groeneveld's version consist of several landclasses (Contributing upland; Ponds; Tarh; Baidh; Lowland), each with the same input data and its own characteristics (areafraction, maximum storage and hydraulic conductivity) which are a result of a calibration, or research done by Cornish, Croke, Kumar, & Karmakar (2012). The original model structure developed by Groeneveld (2012) is shown in Figure 2.
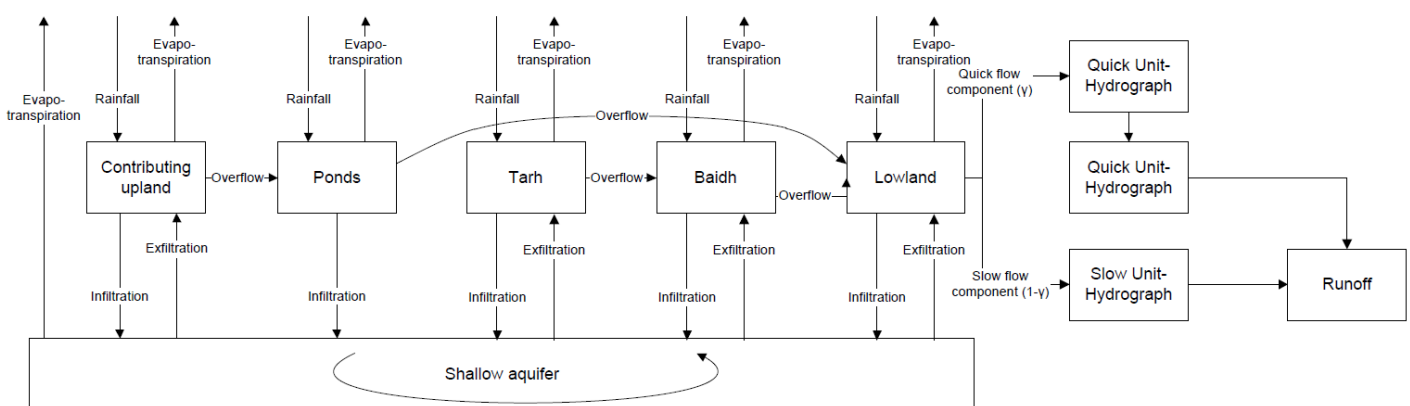

Figure 2 - Original model (Groeneveld, 2012)

The model uses climate data (rainfall and evapotranspiration, 10 minute scale) and the previous storage of the surface to calculate the exfiltration, infiltration, overflow, the new surface storage of the surface module and the evapotranspiration of the shallow aquifer. This calculation follows the sequence of the landclasses shown in Figure 2. After that, the model calculates the quick overflow

runoff component using a unit hydrograph (2 identical linear reservoirs stores in series, i.e. a Nash-cascade) and a slow component runoff using a single hydrograph. These two runoffs together produce the total runoff of the study site at the outlet (Groeneveld, 2012).

Also several options are included in the model. These options are: selecting only small parts of the data for usage, using different catchments with associated calibrated parameters, a scenario analysis and different percentages of pond areas. These options gave the opportunity to predict the effects of several watershed development measures in the modelled study site.

Learning about the model's behaviour required removal of all calibration related model parts, the scenario analyses and the optimization. After that, the parameters are changed back to the initial ones of the start of the research of Groeneveld (2012), which are a result of field work (Cornish, Croke, Kumar, & Karmakar, 2012) . This cleaned model is the start for applying the model to catchments in Andhra Pradesh using a top-down modelling approach, starting with a simple model that is based on field observations (visual interpretation) (Littlewood, Croke, Jakeman, & Sivapalan, 2003).

4. Applying the cleaned model on gauged catchment Andhra Pradesh

This section describes the application of the original model on the gauged catchment in Andhra Pradesh. First a data analysis will be described containing the sub-sections: available data, potential evapotranspiration estimation, auto- and cross correlation analysis and a yearly rainfall – runoff comparison. This section ends with the results of the model run on the gauged catchment.

4.1. Data analysis

4.1.1. Available data

The available stream flow data for this research is the average streamflow per day and is gathered during the period 1 June 1988 – 31 May 1996. This period is therefore the scope of research. The location of the stream flow gauge is displayed in Figure 9.

The available input data for this research consists of daily rainfall data (mm, 0.5 by 0.5 degree grid) and maximum daily temperature data (C, 1 by 1 degree grid), source: Indian Meteorological Department (IMD). These datasets have daily measured data for each grid cell starting in 1971 till the end of 2005, gathered from rainfall gauges within each grid cell. The datasets are adapted for the catchment in Andhra Pradesh by calculating a weighted average based on the area fractions of the catchment matching each grid cell. The rainfall and temperature data of the study period is shown in Figure 3.

Also data from individual rainfall gauges is available. This dataset consist of daily measured rainfall (mm) of each gauge. The dataset has data gaps for some gauges in the catchment, requiring either infilling of the gaps before estimating the areal rainfall, or handling the gaps in the estimation of the areal rainfall.
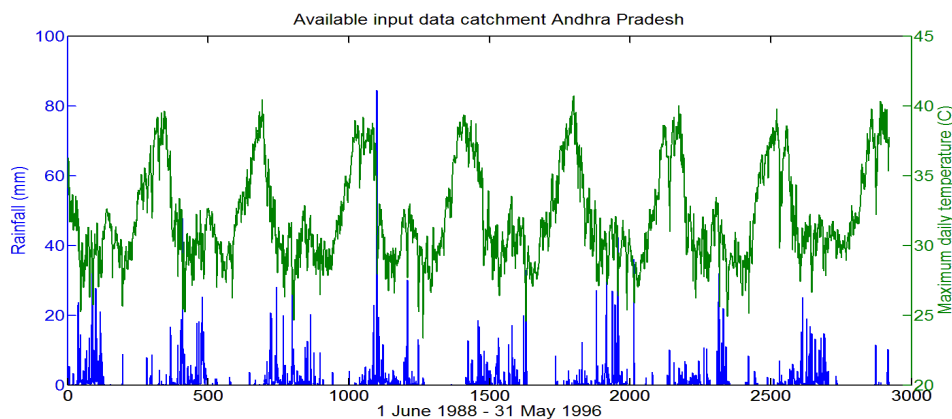

Figure 3 - Available input data (gridded daily rainfall and maximum temperature data)

4.1.2. Estimating potential evapotranspiration

Measured values of evapotranspiration are not available for this area. The only parameter measured in the area that can be used for estimating the evapotranspiration is the maximum daily temperature. Therefore, the evapotranspiration ($\frac{mm}{day}$) is derived from the maximum daily temperature (°C) multiplied by a factor ($\frac{mm}{day*°C}$). For Australia, a common factor calibrated on several datasets is 0.166 (Chapman, 2001). A factor for the Pogro catchment is calculated to use as a representative value for estimating the potential evapotranspiration in the Andhra Pradesh

catchment. The result of this calculation is a multiplier of 0.1446 for the Pogro catchment, further explained in appendix A 1.

4.1.3. Auto- and Cross correlation analysis

Furthermore, an analysis of the cross correlation of daily gridded rainfall data and the observed daily streamflow, as explained in Croke (2005), has been performed. An auto- and cross correlation analysis can provide valuable information about the catchment response to rainfall events in regions where scarce data is available (Bieger, Hormann, & Fohrer, 2012).

The results showed seasonality in the rainfall- and streamflow data, where the streamflow seasonality probably is mostly caused by the rainfall because of the stronger seasonality signal in the rainfall auto correlation, see Figure 4. The expectation is that the model will show comparable seasonality in the modelled flows. In addition, the results showed that the rainfall – streamflow observation does not contain a delay in time series, as can be seen in Figure 5, and therefore the parameters representing the delays (δ) can be set to zero and are not necessary to calibrate further. Also a lot of noise in the data was detected with only a peak correlation coefficient of 0.296, meaning that the gridded data is of poor quality when used for a catchment of 2700 km^2.
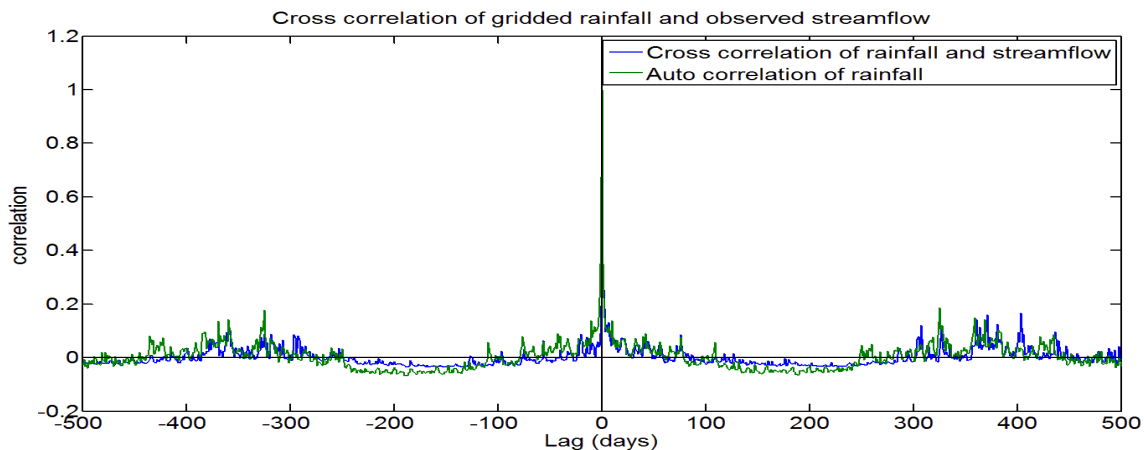

Figure 4 - Cross correlation for Andhra Pradesh catchment using gridded rainfall data, lag between - 500 and + 500 days.
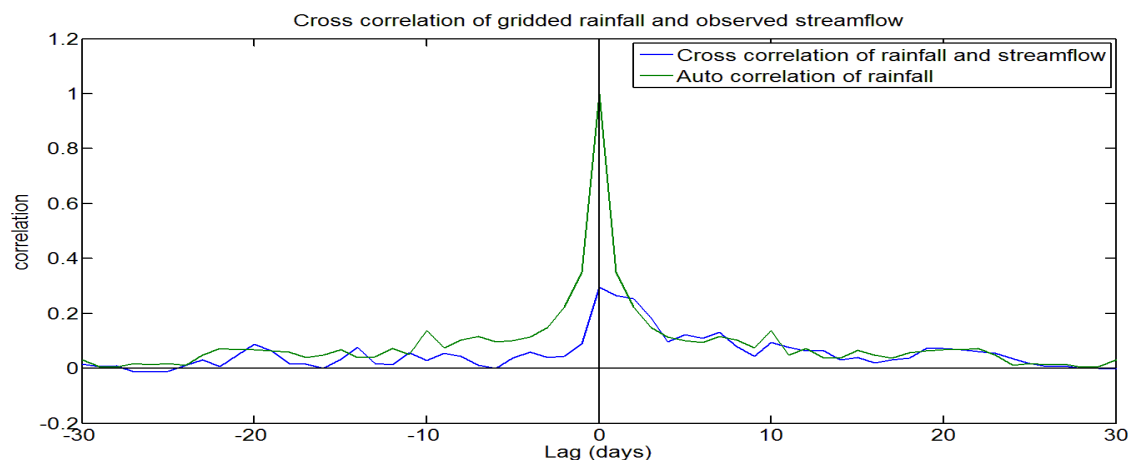

Figure 5 - Cross correlation for Andhra Pradesh catchment using gridded rainfall data, lag between - 30 and + 30 days

Because of the poor quality of the gridded rainfall data and the low correlation coefficient, a cross correlation analysis of averaged individual rainfall gauges in the catchment area has been performed

as well (the rainfall gauges locations are displayed in Figure 9, yellow spots with each gauge denoted by a letter). The straight average of all individual measurements is calculated to use as an estimate for the rainfall in the whole catchment. This data needed some pre-processing before usage was possible. The catchment consists of nine individual rainfall gauges, all gauges recorded daily rainfall measurements in mm. If there was no rainfall event during a particular day, no data has been recorded. This means that the data from all individual gauges only have certain days when rainfall occurs in the database. The rest of the days have been added to this database with a rainfall of zero mm to make the data usable for modelling purposes. Some gauges had a gap in the data, a lack of data in some years. The measurements in the gap (between the two last known measurements) are set to NaN (Not a Number).

The individual rainfall gauges dataset is also analysed using the auto - cross correlation analysis method (Croke, 2005), being showed in Figure 6. The correlation coefficient, after removing a runoff event causing a strange peak in the cross correlation (further explained in appendix A 2) is 0.58 (and the Nash Sutcliffe coefficient added to the model for the calibration afterwards the modification, has increased 0.3 points compared with the gridded rainfall data NS value). This dataset also does not show a delay in time series (Figure 6). Based on the correlation coefficient, the individual rainfall gauges dataset matches the observed streamflow much better in terms of the correlation coefficient and is therefore chosen for use during this research.
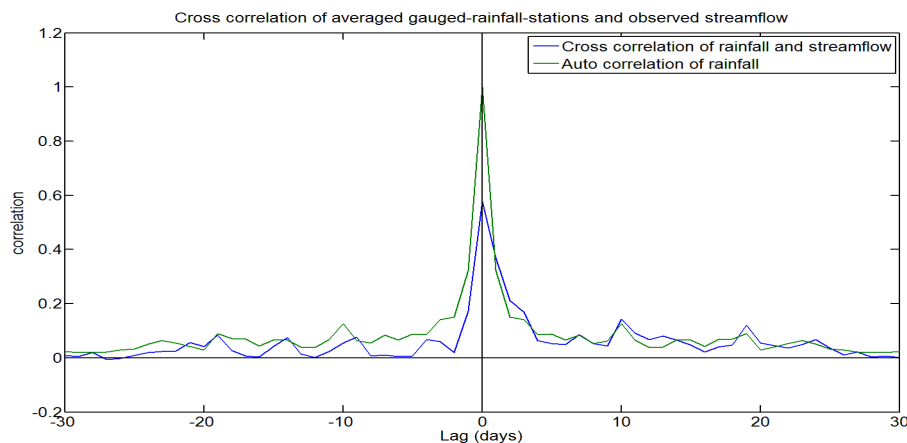

Figure 6 - Cross correlation analysis for Andhra Pradesh catchment, 30 lags. Individual rainfall gauges data

In addition, the first data year has been removed from the performance indicators calculation (the model starts running at the beginning of the dataset to minimize the effects of the setup time on the results). In the first year of data, measurements started on the first of June. This caused a big underestimation of the streamflow in that particular year because it is missing some rainfall events occurred before the first of June, and the model is starting with empty storages. The indicators calculation of the model starts therefore on 1 January 1989 (day number 215).

The year 1990 has been removed because of big observed runoff events without any substantial rainfall occurring in that period. Both rainfall datasets have these problems with year 1990 and therefore this year has been removed from the input data.

4.1.4. Comparison observed yearly rainfall and observed yearly streamflow

After choosing the individual rainfall gauges dataset to continue this research, this dataset has been processed to display yearly rainfall values and compared with yearly observed stream flow values, displayed in Figure 7.
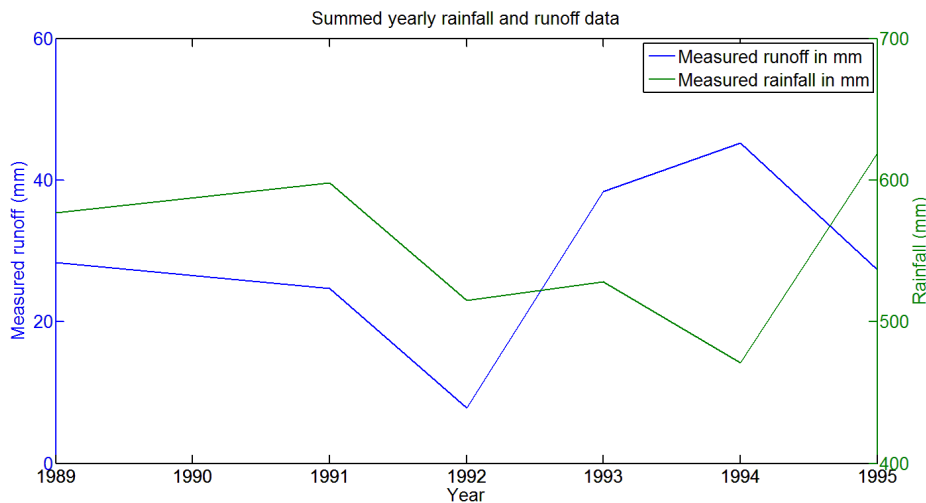

Figure 7 - Yearly observed streamflow and rainfall

As can be seen in Figure 7, the yearly observed runoff is fluctuating significant. This is indicating a change in behaviour in the hydrology of the catchment or an error in the data (e.g. due to a change in the data collection). In 1992 the observed streamflow decreased. However, in 1993 and 1994 the yearly the observed runoff increased without a significant change in the amount of rainfall in those years. Therefore, an underestimation of the modeled streamflow by the model is likely to occur. Because of the decreased obseverd runoff value in 1992 and the increased values in 1993 and 1994, it is more likely that an error in the data occurred instead off a change in model behaviour.

The required dataset and area characteristics extracted from fieldwork to investigate this changed behaviour, or more likely, a change in data collection or an error, is not available at this moment in time. Hence, the exact cause of this difference is outside the scope of this thesis and must be further investigated in additional research to the hydrology of this catchment.

4.2. First model run on gauged catchment Andhra Pradesh

Firstly, the cleaned model was applied to the Andhra Pradesh catchment without adding any structural components to the model. To apply the model on the Andhra Pradesh catchment, the input data must be changed as well the scale of the model. The input is the available daily rainfall data (individual rainfall gauges) and daily maximum temperature data for estimating the evapotranspiration. The catchment area is 2750 km^2, and displayed in Figure 9. The location is displayed in Figure 8, with the catchment area borders shown using a black line.
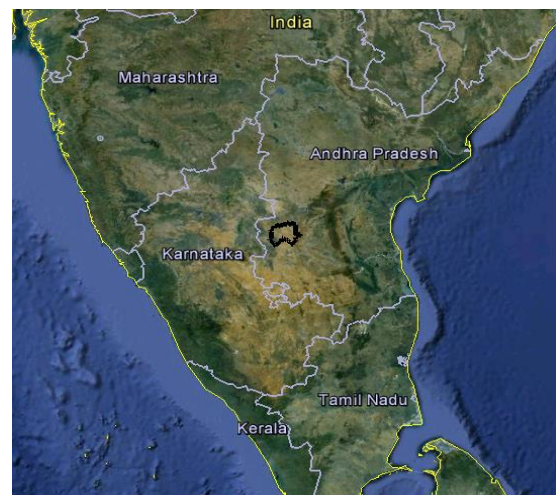

Figure 8 - Location catchment Andhra Pradesh (Google Earth)

A visual inspection of the catchment area using Google Maps / Google Earth (Figure 9) pointed out that the sequence of components being calculated had to be changed. The structure of the model has slightly been changed because of the extension of the contributing area to include the Baidh to fill the ponds, and the absence of an area multiplier to generate more contributing uplands to fill the ponds. The contributing uplands land class has been removed from the model (the uplands are still in the model). The ponds are producing overflow to the lowlands and are being filled by the overflow of the Tarh and Baidh. Google Earth gave an indication of approximately 1 % pond area; 25% Tarh, 40% Baidh and the rest Lowland area. The model run starts in the dry season and with empty stores.

The land classes, as described in the previous paragraph, are located dispersed over the whole catchment area. Exact classifications (e.g. 10 kilometre buffer around the streams are lowlands) of the land classes over the catchment area are not available (only scarce information about the area is available and is mostly derived from related studies, e.g. the WSD study in West Bengal). No information is available to make statements about certain land classes located in particular parts of the catchment area. The only aspect that is validated is the sequence of the land classes (uphill down to the streams) Cornish et al. (2012). Therefore, a very complex model structure including spatial variability is not desirable in this research.
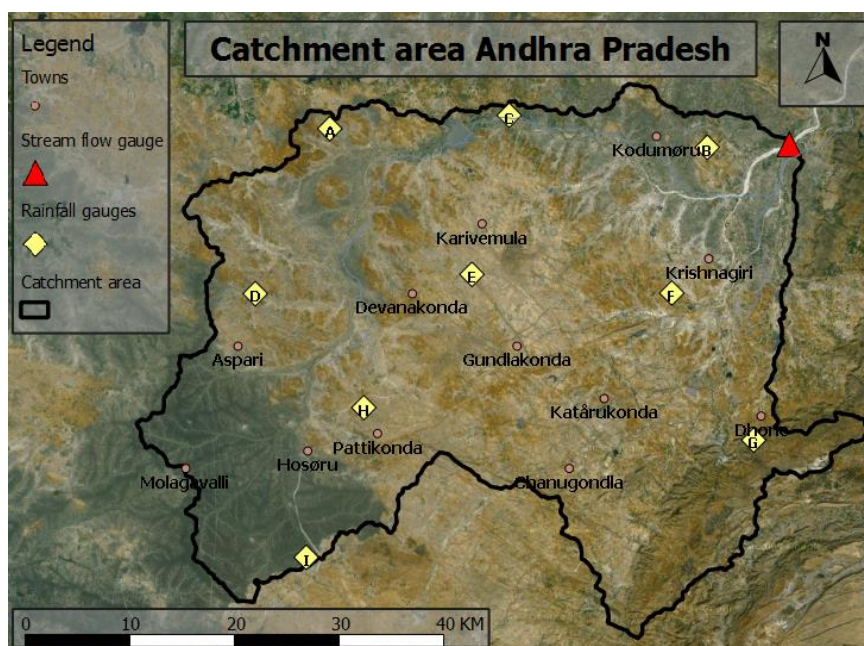

Figure 9 - Catchment Andhra Pradesh

After running the model, the output showed a negative shallow aquifer flow, which is impossible. This resulted in negative water surface storages and evapotranspiration values. The cause of these events was found in the calculation of the evapotranspiration of the shallow aquifer. The evapotranspiration values on occasion exceeded the available storages. A close look to the formulation of the evapotranspiration showed that the formula calculating the evapotranspiration was not formulated well due to an algebraic mistake. A programming formulation that allows the evapotranspiration not to exceed the available shallow aquifer storage plus the infiltration in the shallow aquifer was added to overcome this problem. In section 5.2 this model process will be reformulated.

5. Modification of the original model

This section describes the modifications of the original model applied on the gauged catchment in Andhra Pradesh. Subsection 5.1 describes the modified model structure (with a visual overview). Subsection 5.2 contains the description of the modified model processes, followed by the storages calculation in subsection 5.3. The calibration and the results of the calibration are explained in subsection 5.4. This sections end with the modelled runoff results in subsection 5.5. The final model structure including processes is displayed in Figure 10. The Matlab code is shown in appendix B.

5.1. Model structure

Restructuring the existing model started by looking at the catchment area and the available data. The catchment in Andhra Pradesh consists currently, as mentioned earlier, of Tarh ($_T$), Baidh ($_B$), Lowlands ($_L$) and pond ($_P$) land classes. A comparison of the land use in Andhra Pradesh with the land use in West Bengal revealed several differences. A distinction is made for the catchment in Andhra Pradesh between ponds ($_P$) and dams ($_D$). The ponds are located between the Baidh and lowland land classes. Overflow from the Tarh and Baidh areas will flow in the constructed ponds, which are assumed to have the same characteristics as the constructed ponds in West Bengal.

Because of the spatial magnitude of the gauged catchment, dams are added to the model structure. In the gauged catchment of the Andhra Pradesh, several dams have been built in the river which will hold the water in natural canyons/valleys. This dam land class characterizes the lakes behind the dams. These lakes are much deeper than the ponds and are assumed to have very low infiltration rates because of a rock / non-permeable bottom bed layer. Exfiltration into the lakes behind the dams can occur due to the inflow into the lakes from upstream land units (the exfiltration will occur at the sides of the lake or upstream of the lake).

The lowlands are now located both upstream and downstream of the dams. The difference between these is the overflow of the dams which only occurs downstream. Therefore a distinction is made between those two. The land class Lower Lowlands (LLowlands or $_{LL}$) are the lowlands downstream the dams and the Lowland land class is upstream the dams.

Furthermore, the overflow of the Baidh land class occurs in two different ways in the catchment of Andhra Pradesh. One way is the overflow into the ponds. The other way is that the Baidh produces overflow in the Lower Lowlands too, because the Tarh and Baidh land classes are also located in the catchment area downstream of the dams. A fraction of Baidh land classes is therefore causing overflow directly in the Lower Lowlands. The fraction of the Baidh producing overflow in the ponds (z in the model) is estimated by a visual interpretation of the area. Google Earth indicated a fraction of 25 % (1-z) of the land class Baidh causing direct overflow into the Lower Lowlands.

Also the deep aquifer (subscript $_{da}$) has been added to the model. The catchment of Andhra Pradesh in comparison with the catchment in West Bengal is much dryer, less rainfall (P) occurs (see introduction). The shallow aquifer (subscript $_{sa}$) is much dryer and therefore the inhabitants are pumping water up from the deep aquifer to irrigate their crops. The pumped (PU) water is added to the Baidh storage (S_b) in the model structure because this is the driest cropping land class which will need the water first. The total amount of pumped water from the deep aquifer to the Baidh storage is derived from a survey among villagers in Andhra Pradesh conducted by the LNRMI (Livelihoods and Natural Resources Management Institute) about cropping and water availability issues.

Adding the deep aquifer entails adding the processes between the shallow aquifer and the deep aquifer as well as the storage, or deficit, of the deep aquifer. Water is percolating (F) from the shallow aquifer to the deep aquifer. Because of the dry conditions in the catchment, it is assumed that seepage to the shallow aquifer will not occur and therefore not included in the model.

5.2. Model processes

The new model processes are defined as simple linear functions because no information is available to verify more complex approaches. Exceptions are made when this approach caused problems.

Evapotranspiration

Because of an error in the calculation of the evapotranspiration in the original model, this part has been reformulated. The evaporation of the surface storage (E_s) was formulated as the maximum between the available storage and the potential evapotranspiration (pE) with the residual allocated to the evapotranspiration from the shallow aquifer. Because of the characteristics of the area (semi-arid climate) and the model sequence, the water stored on top of the surface (surface storage, S) that is left after the infiltration (I), exfiltration (EX) and overflow (O) processes, was causing high spiked one day values in the evaporation calculation of the surface layer. Each time step ($_k$) when there was some surface storage left, the evaporation of the surface (E_s) reached the maximum value, draining off the surface storage and in the next time step the evaporation of the surface was zero. This spiked evaporation behaviour is assuming that this area is like a desert where every possible surface storage will directly evaporate, which it is not the case in this catchment.

This problem can be solved in several ways. One solution is to shut off the evaporation of the surface when rainfall occurs, decreasing the potential evaporation because of high air humidity and little sunlight during rainfall. This option resulted in a decreased amount of spikes, but the catchment still acts as a desert climate catchment, as described in the previous paragraph, which is not the case. Another solution is to shut off the evaporation of the surface under wet conditions. This solution assumes that during wet conditions the evaporation will occur through the shallow aquifer evapotranspiration (E_{sa}). This solution needed a change in calculation; first the evapotranspiration of the shallow aquifer is calculated and, after that, the remaining part (pE - E_{sa}) is used for the surface evaporation calculation. This option removed almost all spikes; however, in a semi-arid climate it is not likely that the shallow aquifer will reach such high storage values that are required for this assumption to be valid. The solution was found in calculating an area fraction of the land classes for surface evaporation and assuming that the evapotranspiration of the shallow aquifer occurs first.

The catchment in Andhra Pradesh is a hilly area, so when surface storage occurs, the inundated area is not the same as the total area of that particular land class, see Figure 11. This means that at a certain store level on the surface, S_s, the potential evaporation can only be subtracted from the inundated area fraction. That area fraction (f_s) is a function of the storage available.

The function that indicates the area fraction f_s is assumed as a smooth function between zero and one. Therefore, a Gaussian function is used which has a domain between $-\infty$ and ∞. The left half of the Gaussian function, domain between $-\infty$ and zero, has a smooth increasing function and is a suitable shape for the area fraction calculation. The domain of the cumulative area fraction function is between zero and one. The logarithm of the x values is taken to match this domain with the domain of the Gaussian function.
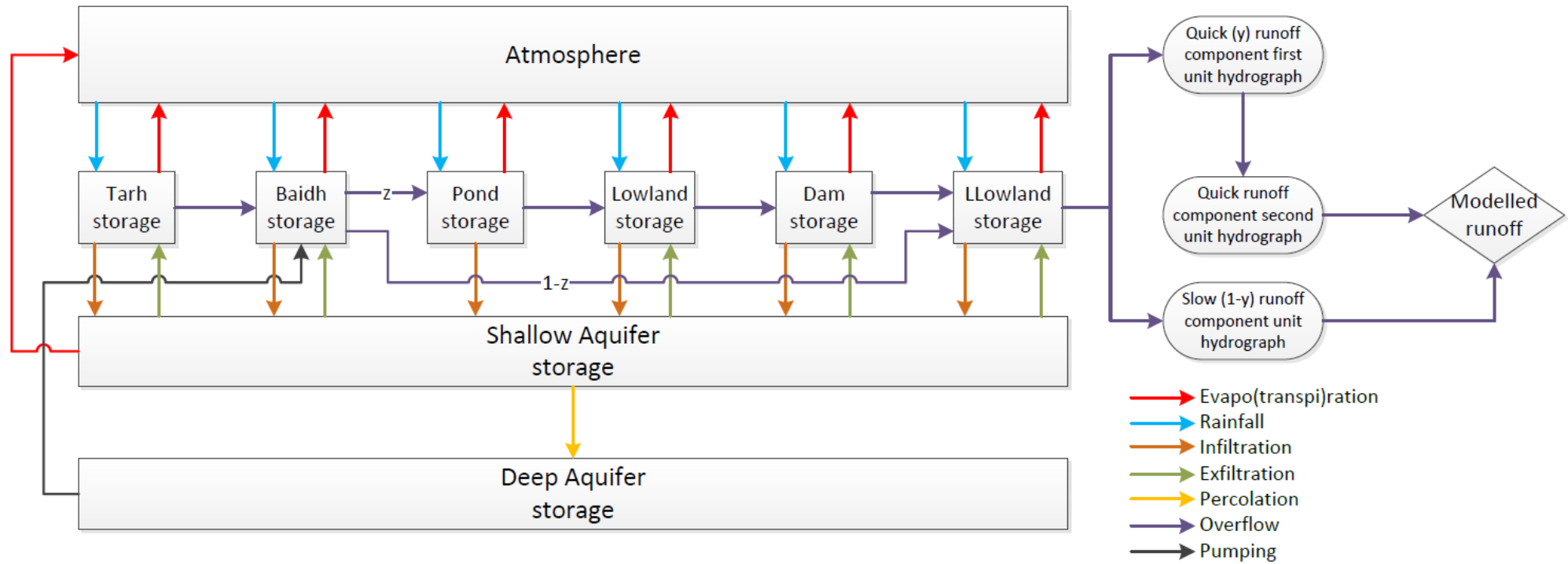
Figure 10 – Modified model structure based the model structure of Groeneveld (2012)
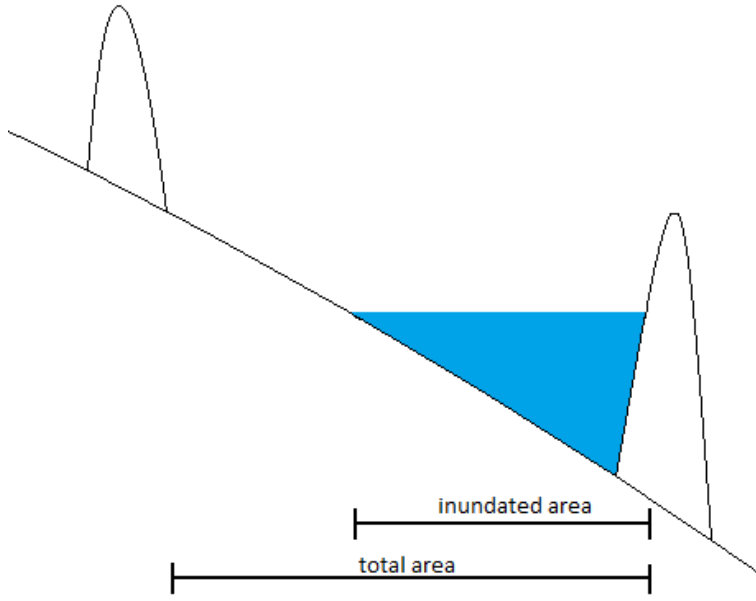
Figure 11 - inundated surface area compared to total area

The formulation of the area fraction can now be derived from the Gaussian function as:

$$f_{s(k)} = e^{-(\frac{x_2}{\mu})^2} = e^{-(\frac{\ln(x_1)}{\ln(\mu)})^2} \tag{1}$$

With $x_2 = \log(x_1)$ and $x_1 = \dfrac{S_{s(k)}}{S_{smax}}$

μ is a parameter that indicates the shape of the cumulative distribution, see Figure 12. Every land class has its own characteristic function dependent on a certain μ value. The μ values give a relative indication of the slope between the areas; an area with a steep slope has a bigger μ value than an area with a lower slope value. Therefore the μ values are set on 0.8; 0.6; 0.5 and 0.5 for Tarh, Baidh, Lowland and Lower Lowland land classes respectively. The ponds in the area are constructed and assumed to have a flat bottom. Hence, this area fraction calculation does not influence the evaporation of the ponds (this is a μ value of one, inundated area fraction of one).
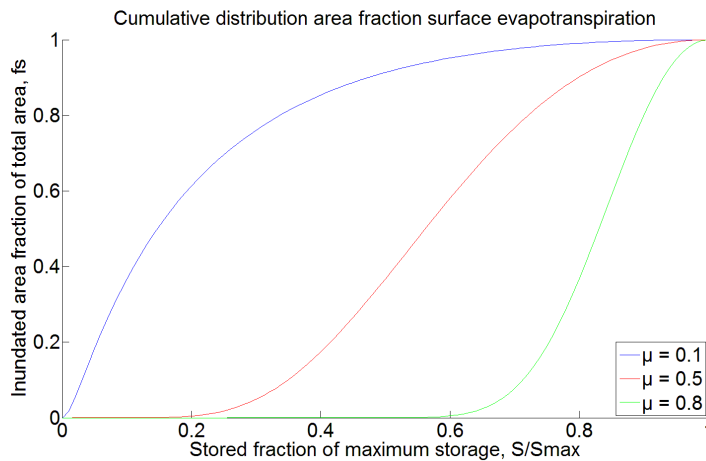

Figure 12 - Cumulative distribution area fraction evapotranspiration

The evaporation of the surface layer can now be calculated as:

$$E_{s(k)} = f_{s(k)} * \max(S_{s(k)}, (pE_{(k)} - E_{sa(k)}))$$ (2)

With $pE_{(k)}$ = maximum daily temperature * 0.1446

The evapotranspiration of the shallow aquifer (E_{sa}) is estimated by a non-linear loss module developed by Croke & Jakeman (2004), first described in Dye & Croke (2003). This non-linear loss module takes the drying effects of the catchment moisture into account. The drying effects are expressed as the Catchment Moisture Deficit (CMD), is this report named as M_{sa} (Moisture deficit shallow aquifer) or M_{da} (Moisture deficit deep aquifer). The relationships between the moisture deficit and the evapotranspiration are defined in Eq. 3 and visualized in Figure 13.

$$E_{sa(k)} = (pE_{(k)}) \text{ for } M_{sa(k)} < h$$

$$= pE_{(k)} * e^{2*(1 - \frac{M_{sa(k)}}{h})} \text{ Otherwise}$$ (3)

Where h is a certain threshold that indicates when the vegetation is beginning to become stressed. When $M_{sa} > h$, the actual shallow aquifer evapotranspiration is assumed to decrease exponentially with increasing M_{sa} (Dye & Croke, 2003; Croke & Jakeman, 2004).
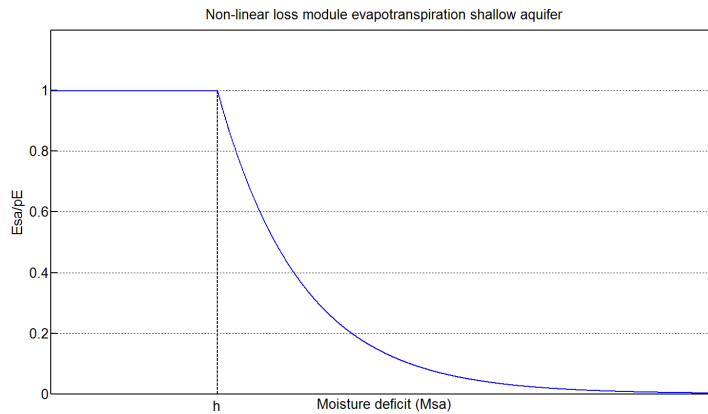

Figure 13 - Evapotranspiration shallow aquifer

Infiltration and overflow

The infiltration processes of the surface storage to the shallow aquifer have not been changed; these calculations following the steps explained in Groeneveld (2012) and are shown in section 5.3. The calculation of the overflow has not been changed either.

Exfiltration

Because of the dry climate and deep groundwater tables in Andhra Pradesh, the exfiltration to the surface is approximately zero in the uphill land classes (Tarh, Baidh and Dams land classes). The calculation of the exfiltration is not based on the shallow aquifer flow but on the moisture deficit of the shallow aquifer. The exfiltration is calculated as a parameter (maximum exfiltration) multiplied by a fraction (e.g. as in Booij, 2012). This calculation is shown in Eq. 4.

$$EX_{(k)} = 0 \text{ for } M_{sa(k)} > g$$

$$= G^* * \frac{g - M_{sa(k)}}{g} \text{ Otherwise} \tag{4}$$

When the moisture deficit has decreased to a certain threshold (g), the exfiltration will increase linearly with the decreasing moisture deficit of the shallow aquifer. G^* is the maximum exfiltration rate. The exfiltration will first occur on the Lower Lowlands and the Lowlands, proceeding uphill.

The exfiltration into the dam land class is calculated due to the impact of the lakes. The lakes caused by the dams have very low infiltration rates (water flowing down into the ground) and are assumed to have circular surface areas. The exfiltration will occur at the sides of the lake, based on the same permeability value as the Lowlands. The area over which the exfiltration into the lakes will occur is the perimeter of the lake multiplied by the depth of the shallow aquifer.

Percolation

The percolation of the shallow aquifer to the deep aquifer is also expressed a fraction multiplied by a parameter that indicates the maximum percolation (F^*) during a time step, see Eq. 5.

$$F_{(k)} = F^* * \frac{f - M_{sa(k)}}{f} \qquad \text{for } M_{sa(k)} < f \text{, otherwise } F_{(k)} = 0 \tag{5}$$

Parameter f is a threshold indicating the value when the percolation process will shut off.

For the percolation and exfiltration calculations these simple approaches were chosen because of the lack of information about the processes. Therefore the parameters will be calibrated in section 5.4. No data is available to verify any more complex approaches for these calculations.

Runoff

The modelled runoff is generated by the overflow of the Lower Lowlands. This overflow is separated in a quick overflow component (γ) and a slow overflow component ($1 - \gamma$). The quick overflow from the Lowlands is using unit hydrographs of two identical linear stores in series (Nash-cascade) producing runoff. The slow overflow component is assumed to be caused by the exfiltration from the shallow aquifer back in the surface storages. This overflow component is translated into runoff using a single unit hydrograph. Both overflow components have a time constant τ and a possible delay δ (which are zero in this research, as pointed out in the data analysis, section 4.1.3) (Groeneveld, 2012), using the following formulas:

$$Q_{(k)} = -\alpha q_{(k-1)} + \beta * U_{(k-\delta)} * (\gamma \text{ or } 1-\gamma) \tag{11}$$

With:
$$\alpha = -e^{-\frac{1}{\tau}}$$
$$\beta = 1 + \alpha$$

5.3. Storages

The surface storage is calculated with a mass balance (as described in e.g. Booij, 2012; Hoekstra, 2012) including the previous defined processes influencing the surface storage, showed in Figure 15. The shallow aquifer storage is defined as the maximum storage minus the deficit at that time step. The deficit is calculated as a mass balance with the processes influencing the shallow aquifer, see Figure 14. The maximum storage is calculated as the depth of the shallow aquifer multiplied by the porosity of the saturated soil. The deep aquifer is only expressed as a deficit with a mass balance.
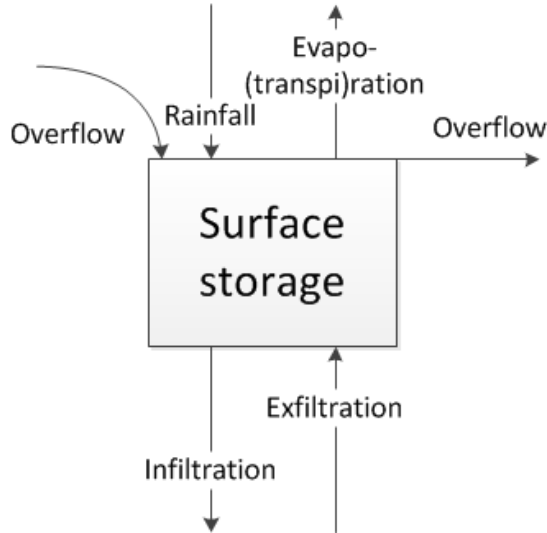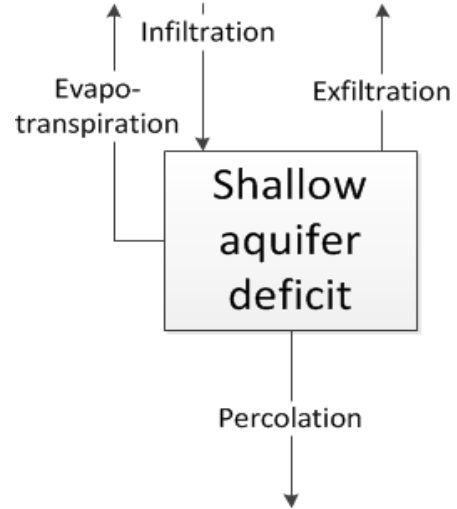
Figure 15 - Surface module

Figure 14 - Shallow aquifer module

Summarized, the formulas to calculate the storages are:

Surface storage of each land class:

$$S_{s(k)} = S_{s(k-1)} + P_{(k)} + I_{(k)} + EX_{(k)} - O_{(k)} - E_{(k)} \tag{6}$$

With:

$EX_{(k)} \quad = G * \dfrac{g - M_{sa(k)}}{g}$ with $M_{sa(k)} < g$, otherwise $EX_{(k)} = 0$

$I_{(k)} \quad = \min(S_{s(k)}, K_{sat} * A_{i(k)})$

$O_{(k)} \quad = \max(S_{(k)} - S_{smax}, 0)$

$E_{(k)} \quad = f_{s(k)} * \max(S_{s(k)}, (pE_{(k)} - E_{sa(k)}))$

Shallow aquifer deficit:

$$M_{sa(k)} = M_{sa(k-1)} - I_{(k)} + EX_{(k)} + E_{sa(k)} + F_{(k)} \tag{7}$$

With:

$EX_{(k)} \quad$ = sum of all exfiltration fluxes to surface storage

$I_{(k)} \quad$ = sum of all infiltration fluxes of the surface storage

$E_{sa(k)} \quad = pE_{(k)} * e^{2*(1 - \frac{M_{sa(k)}}{h})}$ for $M_{sa(k)} > h$, otherwise $E_{sa(k)} = (pE_{(k)})$

$F_{(k)} \quad = F^* * \dfrac{f - M_{sa(k)}}{f}$ for $M_{sa(k)} < f$, otherwise $F_{(k)} = 0$

When the shallow aquifer decifit exceeds the maximum available storage in the shallow aquifer, the exfiltration and the percolation will be zero.

Deep aquifer deficit:

$$M_{da(k)} = M_{da(k-1)} - F_{(k)} \tag{8}$$

5.4. Calibration and validation model on catchment Andhra Pradesh

The model is being calibrated by the Nash-Sutcliffe (NS) coefficient (Nash & Sutcliffe, 1970) and, in addition, the relative volume error (RVE) is used to compare the results. The calibration is done using the 'least square non-linear' (lsqnonlin) function in Matlab. This function solves the given

parameters for the least-squared error in the runoff. The model is calibrated as described in Klemes (1986); first the parameters are calibrated on the first 70% of the dataset and validated on the final 30%. Afterwards, the first 30 % is used for validation and the final 70 % for validation.

The formulas of these calibration indicators are respectively:

$$NS = 1 - \frac{\sum(Q_{mod(k)} - Q_{mes(k)})^2}{\sum(Q_{mes(k)} - avg(Q_{mes}))^2} \tag{9}$$

$$RVE = 100 * \frac{\sum(Q_{mod(k)} - Q_{mes(k)})}{\sum Q_{mes(k)}} \tag{10}$$

With:

$Q_{mod(k)}$ = Modelled runoff on time step k in $\frac{mm}{day}$

$Q_{mes(k)}$ = Measured / observed runoff on time step k in $\frac{mm}{day}$

$avg(Q_{mes})$ = Average off measured / observed runoff of all time steps.

The Nash-Sutcliffe coefficient was used as the performance indicator in the calibration, coupled with a visual interpretation of the modelled stream flow. The advantage of doing a visualisation as pointed out in Bennet, et al. (2012) is that: ''details can be observed in the results which would have remained hidden in a quantitative evaluation, or which can help to direct the tools used for quantitative evaluation. Visualisation takes advantage of the strong human capacity for pattern detection and may allow model acceptance or rejection without determining strict formal criteria in advance'' (Bennet, et al., 2012).

The area characteristics are estimated before the calibration and are displayed in Table 1. The characteristics are derived from a study performed by Cornish et al. (2012). These characteristics will not be changed during the calibration, except for the maximum storage values which can easily be changed by human intervention.

Table 1 - Area characteristics

Parameter values	Tarh	Baidh	Ponds	Lowlands	Dams	LLowlands
Maximum storage (S_{max}, mm)	3	50	3000	100	10000	100
hydraulic conductivity (K, mm/hour)	30	0.5	0.4	0.05	0.005 Infiltration 0.05 Exfiltration	0.05
Area Proportion (A_c)	0.25	0.385	0.005	0.25	0.01	0.10

The parameters that have to be calibrated are:

- Tau quick overflow component (τ_q)
- Tau slow overflow component (τ_s)

- Proportion quick overflow component (γ)
- Threshold MSA percolation (F^*)
- Threshold MSA Exfiltration (G^*)
- Threshold MSA Evapotranspiration (h)
- Exfiltration parameter (g)
- Percolation parameter (f)
- Depth shallow aquifer
- Porosity saturated shallow aquifer
- Initial value moisture deficit shallow aquifer
- Initial value dam storage
- Maximum surface storages $(S_{max\ subscript\ land\ class})$

The quick and slow delay components (δ) are not calibrated because the data analysis pointed out that these parameters are zero, as explained in section 4.1.3.

The calibration on the first 70% of the dataset resulted in a NS and RVE of 0.58 and 13 % respectively and the validation in a NS of 0.61 and RVE of 1.5 %. However, calibration on the final 70% of the dataset resulted in a NS of 0.70 and a RVE of -15.2 % and the validation showed a NS of 0.07 and a RVE of -12.2 %. The calibration and validation of the parameters in two ways according to Klemes (1986) are indicating that a change has been made in either the hydrology of the catchment or an error in the data. The different values of calibration and validation, probably caused by the previous indicated issues, are being underpinned by the results of the yearly rainfall and runoff analysis as performed in section 4.1.4.

The difference in performance indicators is mainly caused by a decreased value of the optimized maximum storage of the baidh landclass during the calibration on the final 70% of the data (from 48 mm to 11 mm) and an increased value of the lower lowland storage (from 23 mm to 51 mm). This resulted in less runoff being generated by the model.

If the calibrated parameters of the first 70% of the dataset are chosen, the ponds will rarely fill. However, Google Earth satellite images showed several filled ponds in the area. Based on the visual interpretation of Google Earth and the expected underestimation of the model as explained in section 4.1.4, the calibrated values of the final 70% of the dataset for the remainder of this research. The performance indicators NS and RVE of the whole dataset with the parameters extracted from the final 70% calibration are 0.66 and -15.8 % respectively, indicating that the hydrology of the catchment is adequately captured, though with a significant bias (RVE).

The results of the calibration are displayed in Table 2 and Table 3. For some runoff events the automatic calibration could not find a proper solution as the calibration gave useless results (i.e. results where outside acceptable ranges, e.g. a negative overflow from the Baidh to the lowland land class meaning that dams do not have any influence on the catchment's outlet). The underestimation of the model can be clarified by the absence of modelling some observed runoff events. This will be explained in the forthcoming part, subsection 5.5.

Table 2 – Calibrated parameter values

Parameter	Value	Unit	Parameter	Value	Unit

Tau quick	0.049	Day	**Exfiltration parameter**	4	mm
Tau slow	12.6	Day	**Percolation parameter**	120	mm
Proportion quick	54.5	%	**Depth shallow aquifer**	10000	mm
Threshold percolation	1725	mm	**Porosity saturated shallow aquifer**	0.4	
Threshold Exfiltration	1725	mm	**Initial value shallow aquifer deficit**	1775	mm
Threshold Evapotranspiration	910	mm	**Initial value dam storage**	9500	mm

Table 3 - Calibrated area characteristics

Parameter values	Tarh	Baidh	Ponds	Lowlands	Dams		LLowlands
Maximum storage (S_{max}, mm)	3	11	3000	126	10000		51
hydraulic conductivity (K, mm/hour)	30	0.5	0.4	0.05	0.005 Infiltration 0.05 Exfiltration		0.05
Area Proportion (A_c)	0.25	0.385	0.005	0.25	0.01		0.10

5.5. Modelled runoff

The output per year being generated by the model is displayed in appendix C. Generally, the model reproduced most runoff events well compared to the observed stream flow. However, a couple of mismatches can be seen in the output. Figure 16 shows the modelled and observed flows in a cumulative graph, indicating where there are significant errors in the modelled flows (day numbers 1400 to 1700 – 1993, and 1900 to 2000 – 1994).
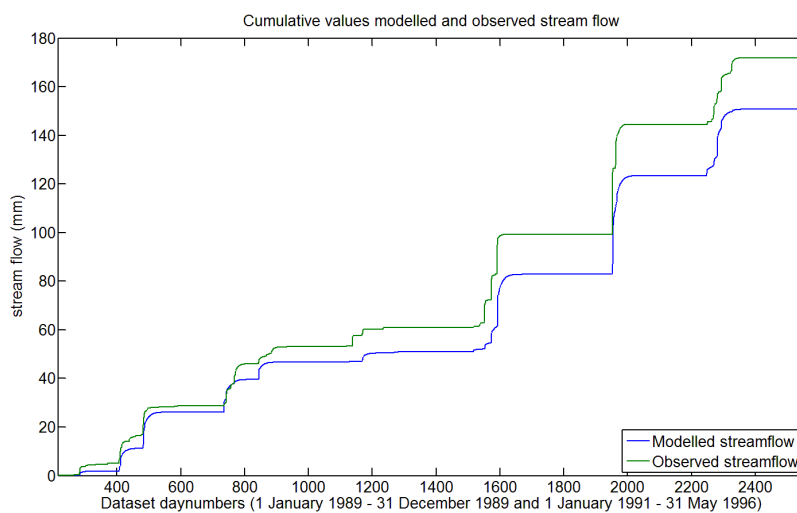

Figure 16 - Cumulative values modelled and observed stream flow

In general, two different possible causes could be identified explaining these modelling differences. The first possibility is spatial variability in rainfall distribution, or errors in the rainfall dataset. The spatial difference of the rainfall can affect the model in overestimating events as well as

underestimating events. The second possible cause is a timing error in one of the datasets. Timing errors are not visible in a cumulative graph.

The underestimation by the model around 31 August 1993 and 21 September 1993 (day numbers 1552 and 1574 respectively), Figure 17, are explaining most of the overall model bias as can be seen in Figure 16. These underestimations are probably caused by the spatial distribution of the rainfall. The rainfall dataset showed that the rainfall events during this period mostly occurred in the North-Eastern part of the catchment (gauges A, B, C and F, Figure 9), near the outlet and a significant lower amount of rainfall occurred in the rest of the area. This rainfall will therefore cause more direct runoff compared to rainfall measured by other gauges in the catchment.
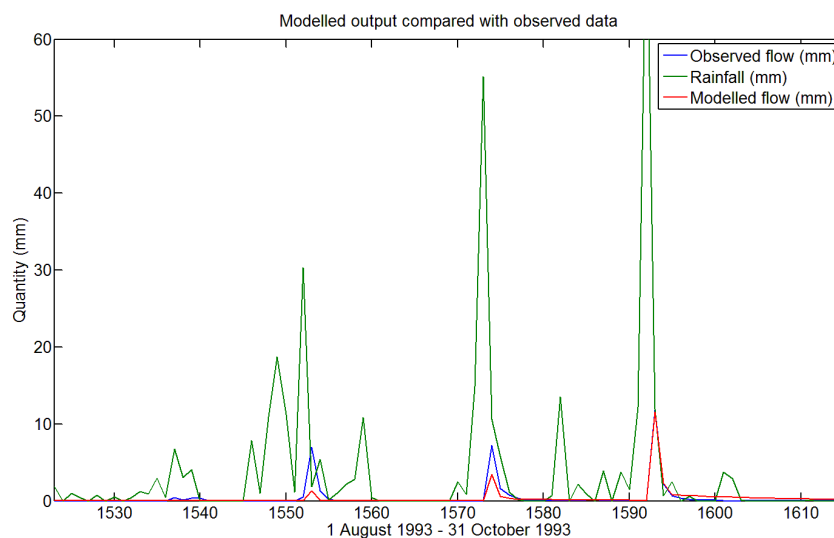

Figure 17 - Output model, selected period 1993

The overestimation in 1995, Figure 18, is probably being caused by the input rainfall data too. This is likely, because of a significant standard error in the mean (20%) of the averaged value of the rainfall at this particular day (day number 2281, 29 August 1995). This statistical error is calculated as the standard deviation of the rainfall gauge data for a particular day, divided by the number of gauges with data for that day.
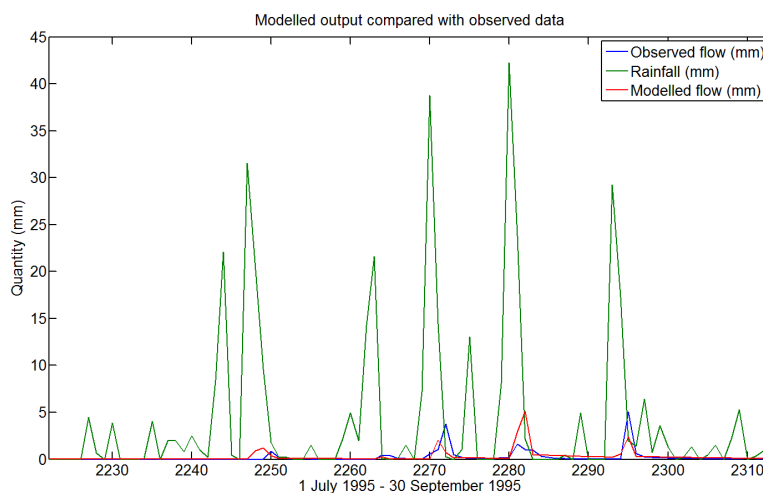

Figure 18 - Output model, selected period 1995

This error is caused by excessive high values of rainfall in a small amount (1 or 2) rainfall gauges. These values could be the real measured values at that point (probably because of small storms causing these amounts of rainfall in a small area), or transcription errors when converting the handwritten databases into digital ones. If these values are the real ones, the calculated average rainfall may be relying too much on these (small) individual events.

The second possible cause, as mentioned before, could be a timing error in one of the datasets, as can be seen in Figure 19. Such errors can easily occur during the data collection or the transformation of the handwritten data to digital databases.
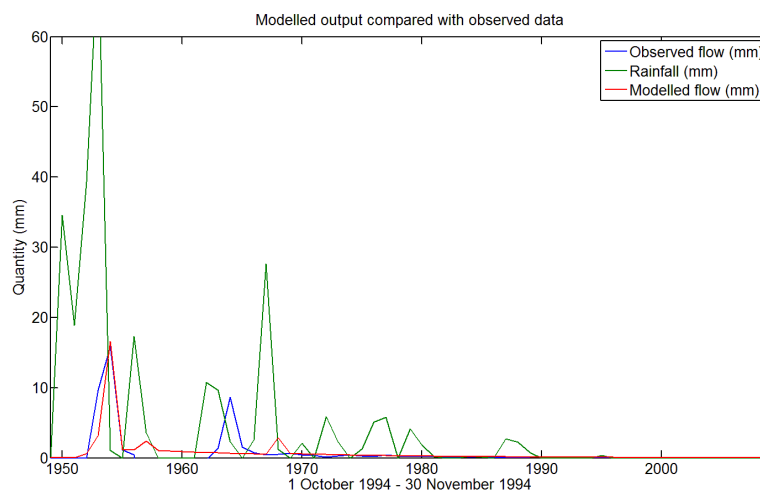

Figure 19 - Output model, selected period 1994

Originally there was an extra peak in the observed stream flow in Figure 19 on day number 1957, 9 October 1994. This peak is removed from the dataset (set to negative values and therefore not used in calculation of performance indicators). The reason for removing this peak was the extra peak in the cross correlation of the rainfall with the observed stream flow data, which was due to this removed stream flow value (see section 4.1.3). This removed value and the mismatch of the modelled and observed stream flow, indicating a possible timing error in streamflow dataset (the underestimated peak must probably be shifted back in time), likely caused by the data collection or transformation.

What the exact cause of these errors is and which dataset is causing these errors, is beyond the limits of this research and should be investigated further during additional research to the hydrology of this catchment.

6. Applying the model on an ungauged study site in Andhra Pradesh

After the original model has been modified to perform as well as possible on the gauged catchment in Andhra Pradesh, the model is applied to an ungauged study site. This study site, called the Gooty study site (called after the local village Gooty), is located South of, and adjacent to the gauged catchment. The Gooty study site is approximately 205 km^2 and shown in Figure 20.
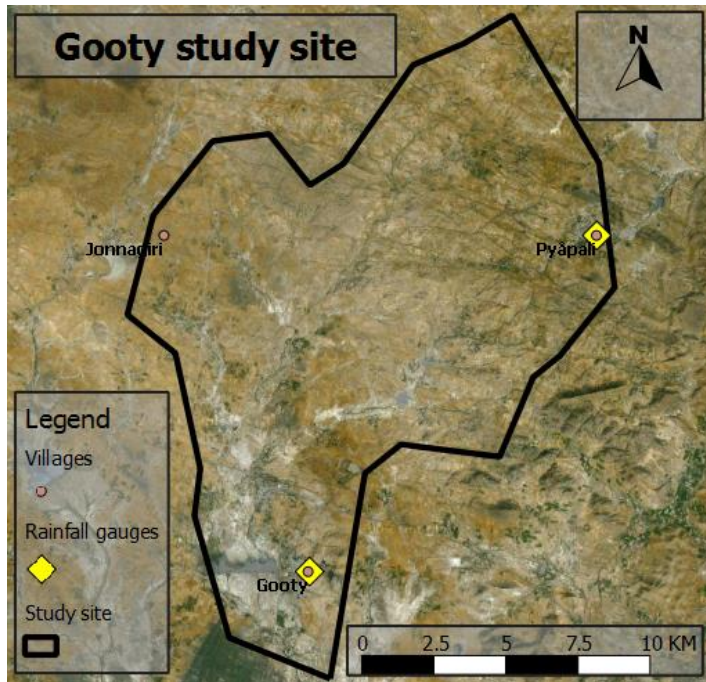

Figure 20 - Gooty study site

To check the model performance on an ungauged study site, a general approach (the calculation of the rainfall-runoff coefficient, RR) has been taken to compare the gauged catchment with the ungauged study site. The rainfall-runoff coefficient is expressed as:

$$RR = \frac{\sum modelled\ runoff}{\sum rainfall}$$ (12)

The rainfall – runoff coefficient of the ungauged study site is calculated by using the same input data sources as the gauged catchment; this means that for the potential evapotranspiration the gridded temperature data multiplied by 0.1446 are used. For the input rainfall data, the individual rainfall gauges in the study site (with the same pre-processing process as described in section 4.1.3) are used. The locations of the individual rainfall gauges are displayed in Figure 20.

For the gauged catchment area in Andhra Pradesh the RR-coefficient of the modified model as described in section 5 is:

$$RR = \frac{151\ mm}{3352\ mm} = 0.045$$

Meaning that 4.5 % of the rainfall will be transferred in runoff by the model (under the assumption that no other in- or out coming fluxes exist in the hydrology of this catchment).

Because of the significant smaller magnitude of the study site and an interpretation of the study site using Google Earth (Figure 20), the dam land class and the lower lowland land class are removed from the model structure. This resulted in a model with the same surface structure as described in Groeneveld (2012) but with the model processes as described in section 5.2 and 5.3 of this research, summarized the model structure is displayed in Figure 21. The Matlab code of this model for the ungauged study site is shown in appendix D.
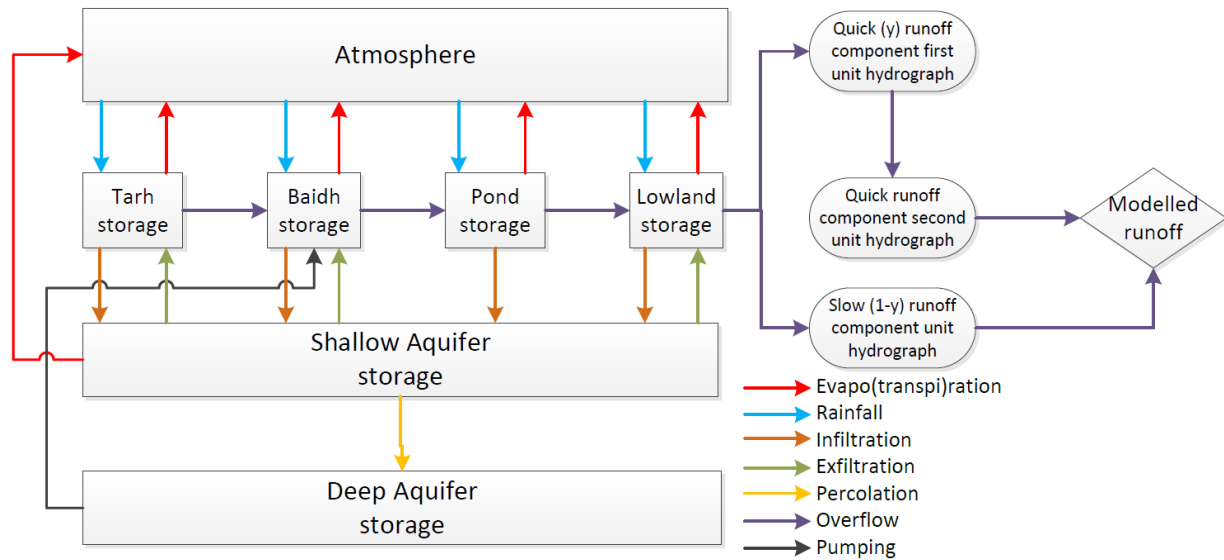

Figure 21 - Model structure Gooty study site

The calibrated parameters, as a result of the previous part of this research (section 5.4), are used to derive the RR-coefficient of the study site. The study site is adjacent to the catchment and therefore assumed to have the same behaviour in hydrology. The RR-coefficient of the Gooty study site as described in this section is:

$$RR = \frac{211\ mm}{3420\ mm} = 0.062$$

The difference with the gauged catchment can be related to the change in rainfall intensity (mm rainfall per event) as well as the change in land use.

The rainfall for the Gooty study site has a higher intensity per event than the rainfall in the gauged catchment, see Table 4. This increased rainfall intensity is causing 95 % more runoff (rainfall dataset applied on the same model, gauged catchment model Andhra Pradesh). The runoff increased from 151 mm total runoff (rainfall dataset catchment Andhra Pradesh) to 295 mm total runoff for the Gooty rainfall dataset.

Table 4 - Rainfall intensity gauged catchment and ungauged study site

Area	Rainfall (mm)	Total rainfall events	Average rainfall intensity (mm/event)
Ungauged Gooty study site	3420 (± 570 mm per year)	364	9.4
Gauged catchment Andhra Pradesh	3353 (± 560 mm per year)	605	5.5

The slightly higher, and much more intense rainfall for the Gooty study site, results in more runoff due to extra overflow. If the rainfall was spread out over a longer period of time, the model would not have generated this amount of rainfall into runoff (evapotranspiration and infiltration would be increased in that case). The increased overflow caused by the rainfall can be related to the input rainfall data for the Gooty study site which only consists of the average 2 measurements. This can led to more extreme values compared to an average of 8 measurements in the catchment area of Andhra Pradesh.

The change in land use in the ungauged study site (removal of the dams and the Lower Lowlands land class), has a significant influence on the modelled runoff too. The runoff decreased to 211 mm (from 295 mm with the same rainfall datasets). This decrease can be explained due to the change in land use. The Lower Lowlands are removed from the model structure which means that the land class producing overflow for the unit hydrographs has a higher maximum storage (from 51 mm Lower Lowlands to 126 mm for the Lowlands), causing less runoff.

The modelled runoff per year being generated for the ungauged Gooty study site is displayed in appendix E.

7. Discussion and conclusion

This section contains the discussion and conclusion. Subsection 7.1 is the discussion about the data analysis and the model results. Subsection 7.2 contains the conclusion of this research followed by a future work section, subsection 7.3.

7.1. Discussion

Data analysis

After a data analysis, which pointed out there is no delay in the data and a streamflow event error had to be removed because of a strange lagged peak in the cross correlation function, the individual rainfall gauges are chosen as the rainfall input data for the remainder of this research. In addition, a factor of 0.1446 has been calculated to extract the potential evapotranspiration from the maximum daily temperature. The yearly observed rainfall-runoff comparison showed a possible change in hydrology, data collection or an error in the data during the years 1992, 1993 and 1994. The observed runoff changed significant during these years without a change in rainfall amounts. The exact cause of this data issue is beyond the limits of this research and is recommended for additional research (see section 7.3).

The results of this research are highly dependent on assumptions being made based on the available scarce information of the catchment and study site, or derived from related studies (e.g. Cornish, et all. 2012) as pointed out in section 4.2. The estimated area fractions of each land class and the spatial variability of the land classes are the major sources of uncertainty in estimating area characteristics. No precise land class classification has been carried out to extract the locations and the area fractions of the land classes, resulting in only an estimation of these characteristics. Spatial variability of land classes is therefore not desirable to include in the model.

Because of the lack of information about the area and the lack of good quality, high resolution and high frequency data, only indications about possible problems with the data could be made. These indications mainly are spatial variability of the rainfall data, timing errors in the datasets and a possible change in behaviour in the hydrology of the catchment or a change in data collection during the data collection period. This resulted in several recommendations for the data and possible additional research themes as summarized in section 7.3.

Model gauged catchment Andhra Pradesh

A visual interpretation of the area using Google Earth resulted in a change of the model structure and the calculation sequence. Also some model processes have been changed because these processes were not included in the original model or were causing problems in the generated output by the model.

The assumptions being made during this research are having great influences on the results generated by the model. Especially the simple linear calibrated percolation and exfiltration processes are influencing the modelled runoff significant. A small change in parameter values will result in largely changed amounts of runoff being generated by the model. Unfortunately, no additional information is available to implement and underpin a more complex approach, or to estimate an order of magnitude. The parameters encapsulating these processes are calibrated during the research, and are the major drivers of uncertainty in the defined model processes.

Further investigation during additional research to this catchment is highly recommendable (see section 7.3).

Nevertheless, the simple structure and processes of the model, based on visual interpretations of the catchment and study site using Google Earth, and applying a simple approach of unknown model processes, gave a better representation of the catchments hydrology than the original model. The Nash-Sutcliffe coefficient and Relative Volume Error of the generated runoff of the model over the whole dataset are 0.66 and -15.8% respectively (using the calibrated parameters of the final 70% of the dataset and the first 30% for validation). The NS coefficient of 0.66 is in agreement with the individual rainfall gauges dataset (higher correlation coeffcient for the individual rainfall gauges resulted in a higher NS coefficient value). The underestimation of the model in relation to the measured streamflow using the chosen parameters, as pointed out in section 4.1.4, was expected because of the increased measured runoff in later years without an increased amount of rainfall in these years. Therefore, this RVE value is within acceptable ranges for model performance evaluation.

Most modelled runoff events are matching the measured runoff. However, some measured runoff events could not be captured by the model. The majority of these mismatches could be related to possible data issues. As a result of these possible data issues, several possible data problems could be identified during the research as pointed out earlier in this section. These identified data problems were able to clarify most modelled errors.

Model ungauged Gooty study site

Applying the modified calibrated version of the model, created during this research, on the ungauged Gooty study site gave comparable results according to the Andhra Pradesh catchments results. The input of the ungauged study site is extracted from rainfall gauges in this area and the gridded temperature data of the gauged catchment. This data is the best available data for this study site.

The visual interpretation of the land use shows that the model structure needed to be changed to suit the study site, the dam and Lower Lowlands land classes have been removed from the model structure. Removal of the dam land class and Lower Lowlands land class in the ungauged study site resulted in a decreased modelled runoff (and a lower rainfall-runoff coefficient). Because of the increased maximum storage value (from 51 mm Lower Lowlands to 126 mm Lowlands) of the land class producing overflow that causes runoff, more rainfall is needed to cause the same amount of runoff.

The calibrated values of the gauged catchment are mainly influencing the exfiltration, infiltration and percolation of the area. Without the dam and lower lowland land classes, these processes would not change significantly and the calibrated parameters therefore are representable for this study site.

The small difference (increase) in rainfall-runoff coefficients of the gauged catchment and the non-gauged study site could mostly be related to a change of input data. The changed input data (rainfall) has more intensive rainfall events, increasing the modelled runoff. The cause of the increased rainfall intensity could also be related to data issues, a scarce of data for this region.

7.2. Conclusion

This research has been performed to apply a model developed for the West Bengal study site to an ungauged study site in Andhra Pradesh by testing the model on a gauged catchment adjacent to the study site. The research started with the original model developed by Groeneveld (2012).

The model developed in this research has approved capability of modelling the catchment and study site in Andhra Pradesh. These models (the code for the catchment and the different code for the study site) can be used for estimating the effects watershed developments in this region of India, trying to improve the livelihood of farmers.

A very complex model structure and model processes including spatial variability could have been chosen to perform this research. However, no information is available to underpin complex approaches. Therefore, a model structure and processes defined as simple as possible are chosen for performing this research.

7.3. Future work

As pointed out several times during the research and summarized in the discussion and conclusion section, most opportunities to improve this model and other model purposes for this region in India are in the improvement of the data; gathering more frequently reliable data with a higher resolution and in the investigation of area characteristics and unknown model processes.

During the research, limited information about the area was available. This caused large sources of uncertainty in the assumptions of the area characteristics and model processes. More research should be done to the area characteristics and the hydrology of this catchment and study site to decrease these uncertainty sources. Especially the exfiltration and percolation processes, which are now simple linear calibrated processes, require additional research to get an improved representation of the hydrology in the area. Also an area classification of the land classes would improve the representation of the area by a model due to the opportunity to include spatial variability into the model.

During the research some possible data issue problems were identified. The errors that led to the identification of these data issues should be investigated further, trying to extract the exact cause of the problems and if these are related to data problems, make recommendations about how to improve the available data and data quality. This improved data can be used in forthcoming research to this region for many purposes, hydrological modelling purposes in particular. Especially more frequent measurements of streamflow and rainfall (e.g. hourly or a 10 minute scale) and an increase of the resolution of the rainfall data can provide valuable inputs for forthcoming research (will decrease extreme values and give input for spatial variability in the rainfall).

Another additional research theme recommended, concerns the change in the runoff coefficient towards the end of the dataset, due to either changed hydrological behaviour or to artefacts in the data of the gauged catchment in Andhra Pradesh. The calibration and validation according to Klemes (1986) and the data analysis indicated a change in behaviour, an error in the data, or a change in data collection after the wet season of 1992. What is causing this change in measured runoff is beyond the limits of this research and should be investigated further to get a better understanding of possible changes in the catchment.

8. Bibliography

ACIAR. (2003, March 13). Retrieved June 14, 2013, from Australian Centre for International Agricultural Research: aciar.gove.au/aboutus

Bennet, N. D., Croke, B. F., Guariso, G., Guillaume, J. H., Hamilton, S. H., Jakeman, A. J., et al. (2012). Characterising performance of environmental models. *Environmental Modelling & Software, Volume 40, Issue 1, 25 September*, Pages 1 - 20. DOI: 10.1016/j.envsoft.2012.09.011.

Bieger, K., Hormann, G., & Fohrer, N. (2012). Using residual analysis, auto- and cross-correlations to identify key processes for the calibration of the SWAT model in a data scarce region. *Advances in Geosciences*, Pages 23 - 30. DOI 10.5191/adgeo-31-23-2012.

Booij, M. J. (2012). *Inleiding Waterbeheer.* Enschede: University of Twente.

Chapman, T. G. (2001). Estimation of Daily Potential Evaporation for Input to Rainfall-Runoff Models. *MODSIM2001: Integrating Models for Natural Resources Management Across Disciplines, Issues and Scales, Vol 1. MSSANZ*, (pp. 293 - 298).

Cornish, P. S., Kumar, A., & Khan, M. (2010, November). *East India Plateau - Basket Case, or Future Food Basket*. Retrieved June 14, 2013, from Australian Society of Agronomy: http://www.regional.org.au/au/asa/2010/farming-systems/international/6957_cornishp.htm

Cornish, P., Croke, B. F., Kumar, S., & Karmakar, D. (2012). *Water harvesting and better cropping systems for smallholders of the East India Plateau.* ACIAR: Canberra.

Croke, B. F. (2005). Land use impacts on hydrologic response in the Mae Chaem catchment, Northern Thailand. *International Conference on Simulation and Modelling.* Nakornpathom: SimMod.

Croke, B. F., & Jakeman, A. J. (2004). A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environmental Modelling & Software, Volume 19, Issue 1*, Pages: 1 -5 DOI: 10.1016/j.envsoft.2003.09.001.

Croke, B. F., & Jakeman, A. J. (2005). Corrigendum to '' A Catchment Moisture Deficit module for the IHACRES rainfall-runoff model''. *Environmental Modelling & Software, Volume 20*, Page 977. DOI: 10.1016/j.envsoft.2004.11.004.

Dye, P. J., & Croke, B. F. (2003). Evaluation of streamflow predictions by the IHACRES rainfall-runoff model in two South African catchments. *Environmental Modelling & Software, Volume 18*, Pages 705 - 712. DOI:10.1016/S1364-8152(03)00072-0.

Gillet, A., Hammond, A., & Martala, M. (2009). *Succesful academic writing.* Harlow: Pearson Education Limited.

Government of Andhra Pradesh. (2011). *State profile-Andhra pradesh.* Retrieved 3 26, 2013, from AP state portal: http://www.ap.gov.in/Other%20Docs/COMPARATIVE%20STATISTICS%20%28STATES%29.pdf

Groeneveld, D. H. (2012). *Modelling watershed development in West Bengal (East India Plateau) using a small scale rainfall-evapotranspiration-runoff model.* unpublished.

Hoekstra, A. Y. (2012). *Water.* Enschede: University of Twente.

Jakeman, A. J. (2013, 9 7). *Integrated Catchment Assessment and Management (iCAM).* Retrieved 3 22, 2013, from Ferner school of Envrinment & Society; ANU College of Medicine, Biology & Environment: http://fennerschool-research.anu.edu.au/icam/

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

Littlewood, I. G., Croke, B. F., Jakeman, A. J., & Sivapalan, M. (2003). The role of 'top-down' modelling for Prediction in Ungauged Basins (PUB). *Hydrological Processes, Volume 17*, Pages 1673 - 1679. DOI: 10.1002/hyp.5129.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual modles part 1 - A discussion of principles. *Journal of Hydrology, Volume 10, Issue 3*, Pages 282 - 290. DOI: 10.1016/0022-1694(70)90255-6.

Syme, G. (2009). *Impacts of meso-scale Watershed Development in Andhra Pradesh (India) and their implications for designing and implementing improved WSD policies and programs*. Retrieved 3 24, 2013, from ACIAR; Australian Centre for International Agricultural Research: http://aciar.gov.au/project/LWR/2006/072

World Weather Online. (2013). *Kurnool Weather, India Weather averages*. Retrieved 5 30, 2013, from World Weather Online: http://www.worldweatheronline.com/Kurnool-weather-averages/Andhra-Pradesh/IN.aspx

Appendices

A. Data analysis

1. Factor calculation for estimating potential evapotranspiration

To estimate the potential evapotranspiration (PET) for the catchment in Andhra Pradesh, the Pogro catchment is used to derive a conversion factor between PET and maximum daily temperature. The available data of the Pogro catchment was converted to daily data to get a factor that represents calculation on daily basis, like the data for the catchment in Andhra Pradesh. This comparison resulted in a factor 0.1446 for the Pogro catchment. This factor is assumed to be a good representative for the catchment in Andhra Pradesh.

The result of the calculation (0.1446) compared with the factor calibrated by Chapman (2001) for Australia, means that the potential evapotranspiration at a certain temperature is lower in West Bengal compared to Australia. This difference in multiplier values can be explained by the fact that in the dry season of India (periods with high temperatures and no rainfall) evapotranspiration is significantly reduced because of the absence of surface storage and sub-surface storage. In addition, during the beginning of the wet season, the air humidity is high, which results in the potential evapotranspiration being approximately zero. As most of the evapotranspiration occurs during or immediately after the wet season, the higher PET in the dry season is not significant in terms of the application to a rainfall-runoff model.

2. Removed peak observed streamflow data

In the cross correlation analysis of the rainfall and streamflow data a strange peak occurs at a three-day lag (Figure 22). This indicates a possible timing error in the streamflow data. The dataset was searched for peaks in the data occurring after three days. One peak, occurring around 10 October 1994, was removed from the streamflow dataset (set to negative values and not used in the calculation of the cross correlation function). This resulted in an increased correlation coefficient to 0.58. As can be seen in the cross correlation graph of the dataset with the removed peak, Figure 6 – section 4.1.3, the peak at lag = 3 has almost totally disappeared and therefore that peak was mostly caused by the runoff event on 10 October 1994.
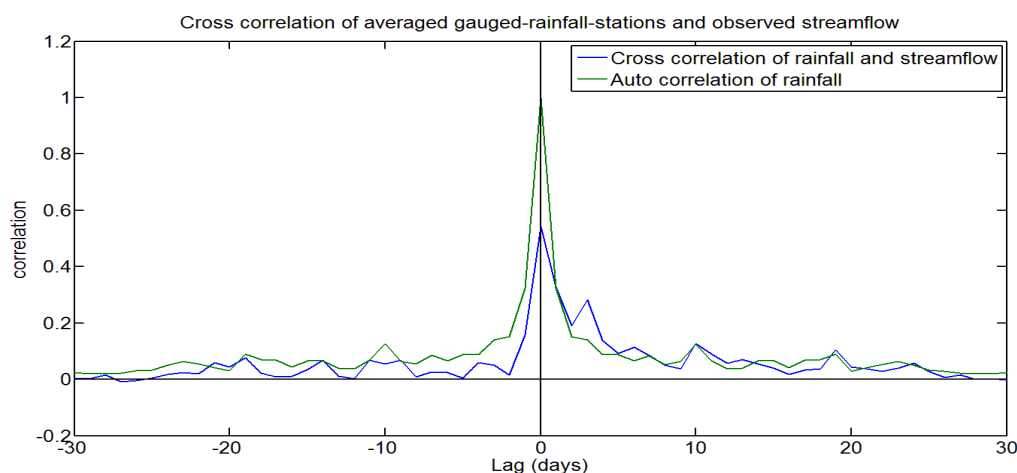

Figure 22 – Three-day lag-peak cross correlation

B. Matlab code modified model for gauged catchment Andhra Pradesh

The Matlab code of the end of phase 2 (applying the model on a gauged catchment in Andhra Pradesh) is shown here. This code is the base for phase 3 of this research.

```matlab
% Model Andhra Pradesh gauged catchment
% Modified version, Final version by Rinse Wilmink
% r.j.a.wilmink@student.utwente.nl

clear all
close all
clc

%---- Input Data ----------------------------------------------
load pq_AP.csv
load temp_AP.csv
load P_new_AP.csv
Data1 = pq_AP;
Data2 = temp_AP;
Data3 = P_new_AP;

%Data 1: rainfall,flow,day,days year, year,P cel 1,P cel 2,P cel 3,P cel 4,
%        day,month, flow with negative values, flow with zero values (in
%        steat of negative ones)
%        (mm)    ,m3/s,nr, 365     , nr , mm    , mm    , mm    , mm,
%        nr,nr,m3/s,m3/s

%Data 2: temp, day,days year,year,temp c1,temp c2, temp c3, temp c4
%        C   , nr , 365     , nr , C      , C      , C       , C

%Data 3: 9 coloms rainfall individual rainfall gauges, colom 10 is average
%rainfall (mm/day), colom 11 = date

x = [0.0491 12.6  0   0 54.5 1725  1725 910  4   120  10000  0.4 0.75] %#ok<NOPTS>
%    x1      x2   x3  x4 x5   x6    x7   x8  x9  x10  x11    x12 x13

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% x1 Tau quick
% x2 Tau slow
% x3 Delay quick
% x4 Delay slow
% x5 Proportion quick (100%)
% x6 Threshold MSA F
% x7 Threshold MSA EX (g in report)
% x8 Threshold MSA eP SA (h in report)
% x9 Exfiltration parameter (mm/dag) (G in report)
%x10 Percolation parameter (mm/dag)  (F in report)
%x11 depth shallow aquifer (mm)
%x12 porosity saturated shallow aquifer
%x13 Fraction Overflow dams z
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% t1 and t2 for period selection (t1 to t2) graphs
% Whole data set: set to: k = 2:size(Data,1)
t1 = 215;     % originally t1 = 2 (start dataset), t 215 = 1 January 1989
t2 = 2556;    % originally t2 = 2557-1 ( end dataset)
n = 1;        % startpoint FOR-loop and initial values MSA, MDA

% input: rainfall, pet
Zeros = zeros(size(Data1,1),1); %starting empty

flow = Zeros;
temp = Zeros;
flow(:,1) = Data1(:,12);
temp(:,1) = Data2(:,1);
P = Data3(:,10);
pE = temp.*0.1446;
Q = flow;
Q(Q == -99) = NaN;

timestep = 24;      % timestep in days, 24 hour
A = 2750 .* 10^6;   % Area A is 2.750 sqkm

% Surface Module Parameters

%%% Tarh (T)
T_Smax = 3;      % 3 originally;
T_Ksat = 30;
T_Ac = 0.25;
%%% Baidh (B)
B_Smax = 11;     % 50 originally;
B_Ksat = 0.5;
B_Ac = 0.385;
%%% Ponds (P)
P_Smax = 3000;   % 3000 originally;
```

```matlab
P_Ksat = 0.4;
P_Ac = 0.005;
%%% Lowland (L)
L_Smax = 126;       % 100 originally;
L_Ksat = 0.05;
L_Ac = 0.25;
%%% Dams (D)
D_Smax = 10000;
D_Ksat = 0.005;
D_EXKsat = 0.05;
D_Ac = 0.01;
%%% LLowland under dams (LL)
LL_Smax = 51;       % 100 originally;
LL_Ksat = 0.05;
LL_Ac = 0.10;

MSA = Zeros;        % Moisture deficit shallow aquifer
MSA(n) = 1775;      % Start value
MDA = Zeros;        % Moisture deficit deep aquifer
MDA(n) = 2000;      % Start value

%%% Discharge q parameters
tauq = (x(1));      %(day)
taus = (x(2));      %(day)

alphaq = -exp(-1./tauq);
betaq = 1+alphaq;

alphas = -exp(-1./taus);
betas = 1+alphas;

T_Ksat = T_Ksat .* timestep;
B_Ksat = B_Ksat .* timestep;
P_Ksat = P_Ksat .* timestep;
L_Ksat = L_Ksat .* timestep;
D_Ksat = D_Ksat .* timestep;
LL_Ksat = LL_Ksat .* timestep;

T_S = Zeros; %starting empty
T_O = Zeros;
T_E = Zeros;
T_I = Zeros;
B_S = Zeros; %starting empty
B_O = Zeros;
B_E = Zeros;
B_I = Zeros;
P_S = Zeros; %starting empty
P_O = Zeros;
P_E = Zeros;
P_I = Zeros;
L_S = Zeros; %starting empty
L_O = Zeros;
L_E = Zeros;
L_I = Zeros;
D_S = Zeros; %starting empty
D_S(n) = 9500; %inital value dam storage
D_O = Zeros;
D_E = Zeros;
D_I = Zeros;
LL_S = Zeros; %starting empty
LL_O = Zeros;
LL_E = Zeros;
LL_I = Zeros;
F = Zeros;          % Peroclation from shallow aquifer
PU = Zeros;     % Pumping

% Area fractions surface evapotranspiration
T_fs = Zeros;
B_fs = Zeros;
L_fs = Zeros;
LL_fs = Zeros;

I = zeros(length(Zeros),8);
E = zeros(length(Zeros),9);
O = zeros(length(Zeros),8);
S = zeros(length(Zeros),7);
EX = zeros(length(Zeros),9);
q = zeros(length(Zeros),6);

LL_AcE = 0;
D_AcE = 0;
L_AcE = 0;
B_AcE = 0;
T_AcE = 0;

%%% Calculation

for k = (n+1):size(Data1(:,1)) % whole dataset
```

```matlab
%%% Exfiltration
if MSA(k-1) < x(7);
    EX(k,2) = x(9).*((x(7) - MSA(k-1))./x(7));
else
    EX(k,2) = 0;
end
EX(k,3) = min(EX(k,2),((LL_Ksat.*LL_Ac)+(L_Ksat.*L_Ac)));
EX(k,4) = EX(k,3).*LL_Ac./(LL_Ac+L_Ac);
EX(k,5) = EX(k,3).*L_Ac./(LL_Ac+L_Ac);

if EX(k,2)-EX(k,3) > 0
    EX(k,6) = min(EX(k,2)-EX(k,3),D_EXKsat.*(x(11).*2.*pi.*sqrt(D_Ac./pi)));
    if EX(k,2)-(EX(k,3)+EX(k,6)) > 0
        EX(k,7) = min(EX(k,2)-(EX(k,3)+EX(k,6)+...
            EX(k,7)),B_Ksat.*B_Ac);
        if EX(k,2)-(EX(k,3)+ EX(k,6)+EX(k,7)) > 0
            EX(k,8) = min(EX(k,2)-(EX(k,3)+EX(k,6)+...
                EX(k,7)+EX(k,8)),T_Ksat.*T_Ac);
        end
    end

end
LL_AcE = EX(k,4)./(LL_Ksat.*LL_Ac);
D_AcE = EX(k,6)./(D_Ksat.*D_Ac);
L_AcE = EX(k,5)./(L_Ksat.*L_Ac);
B_AcE = EX(k,7)./(B_Ksat.*B_Ac);
T_AcE = EX(k,8)./(T_Ksat.*T_Ac);

%%% Calculating evaportanspiration shallow aquifer
E(k,8) = (pE(k-1)).*min(1,exp(2.*(1-MSA(k-1)/x(8))));

%%% Pumping from deep aquifer
PU_days = [0:45,280:365,610:695,1005:1090,1380:1445,1745:1830,2135:220];
if ismember(k,PU_days) == 1;
    PU(k) = 0.262;
else
    PU(k) = 0.0;
end

%%% Tarh (T)
T_S(k) = T_S(k-1) + P(k-1);
T_I(k) = min(T_S(k),T_Ksat.*(1-T_AcE));
T_S(k) = T_S(k) - T_I(k) + EX(k,8)./T_Ac;
T_O(k) = max(T_S(k)-T_Smax,0);
T_S(k) = T_S(k) - T_O(k);

T_u = 0.8;
T_fs(k) = exp(-(log(T_S(k)./T_Smax)./log(T_u)).^2);

T_E(k) = min(T_S(k),T_fs(k).*(pE(k-1)-E(k,8)));
T_S(k) = T_S(k) - T_E(k);

%%% Baidh (B)
B_S(k) = B_S(k-1) + P(k-1) + T_O(k).*T_Ac./B_Ac + PU(k)./B_Ac;
B_I(k) = min(B_S(k),B_Ksat.*(1-B_AcE));
B_S(k) = B_S(k) - B_I(k) + EX(k,7)./B_Ac;
B_O(k) = max(B_S(k)-B_Smax,0);
B_S(k) = B_S(k) - B_O(k);

B_u = 0.6;
B_fs(k) = exp(-(log(B_S(k)./B_Smax)./log(B_u)).^2);

B_E(k) = min(B_S(k),B_fs(k).*(pE(k-1)-E(k,8)));
B_S(k) = B_S(k) - B_E(k);

%%% Ponds (P)
P_S(k) = P_S(k-1) + P(k-1) + x(13).*B_O(k).*B_Ac./P_Ac;
P_I(k) = min(P_S(k),P_Ksat);
P_S(k) = P_S(k) - P_I(k);
P_O(k) = max(P_S(k)-P_Smax,0);
P_S(k) = P_S(k) - P_O(k);
P_E(k) = min(P_S(k),(pE(k-1)-E(k,8)));
P_S(k) = P_S(k) - P_E(k);

%%% Lowlands (L)
L_S(k) = L_S(k-1) + P(k-1) + P_O(k).*P_Ac./L_Ac;
L_I(k) = min(L_S(k),L_Ksat.*(1-L_AcE));
L_S(k) = L_S(k) - L_I(k) + EX(k,5)./L_Ac;
L_O(k) = max(L_S(k)-L_Smax,0);
L_S(k) = L_S(k) - L_O(k);

L_u = 0.5;
L_fs(k) = exp(-(log(L_S(k)./L_Smax)./log(L_u)).^2);

L_E(k) = min(L_S(k),L_fs(k).*(pE(k-1)-E(k,8)));
L_S(k) = L_S(k) - L_E(k);
```

```matlab
%%% Dams (D)
D_S(k) = D_S(k-1) + P(k-1) + L_O(k).*L_Ac./D_Ac;
D_I(k) = min(D_S(k),D_Ksat.*D_Ac);
D_S(k) = D_S(k) - D_I(k) + EX(k,6);
D_O(k) = max(D_S(k)-D_Smax,0);
D_S(k) = D_S(k) - D_O(k);
D_E(k) = min(D_S(k),(pE(k-1)-E(k,8)));
D_S(k) = D_S(k) - D_E(k);

%%% Lower Lowlands (LL)
LL_S(k) = LL_S(k-1) + P(k-1) + D_O(k).*D_Ac./LL_Ac + (1-x(13)).*B_O(k).*B_Ac./LL_Ac;
LL_I(k) = min(LL_S(k),LL_Ksat.*(1-LL_AcE));
LL_S(k) = LL_S(k) - LL_I(k) + EX(k,4)./LL_Ac;
LL_O(k) = max(LL_S(k)-LL_Smax,0);
LL_S(k) = LL_S(k) - LL_O(k);

LL_fs(k) = exp(-(log(LL_S(k)./LL_Smax)./log(L_u)).^2);

LL_E(k) = min(LL_S(k),LL_fs(k).*(pE(k-1)-E(k,8)));
LL_S(k) = LL_S(k) - LL_E(k);

%%% Surface module output Infiltration, Evaporation, Overflow, Storage
I(k,2) = T_I(k).*T_Ac + B_I(k).*B_Ac + P_I(k).*P_Ac + L_I(k).*L_Ac + D_I(k).*D_Ac + LL_I(k).*LL_Ac;
I(k,3) = T_I(k);
I(k,4) = B_I(k);
I(k,5) = P_I(k);
I(k,6) = L_I(k);
I(k,7) = D_I(k);
I(k,8) = LL_I(k);

E(k,2) = T_E(k).*T_Ac + B_E(k).*B_Ac + P_E(k).*P_Ac + L_E(k).*L_Ac + D_E(k).*D_Ac + LL_E(k).*LL_Ac;
E(k,3) = T_E(k);
E(k,4) = B_E(k);
E(k,5) = P_E(k);
E(k,6) = L_E(k);
E(k,7) = D_E(k);
%E(k,8) = shallow aquifer evapotranspiration
E(k,9) = LL_E(k);

O(k,2) = LL_O(k).*LL_Ac;
O(k,3) = T_O(k);
O(k,4) = B_O(k);
O(k,5) = P_O(k);
O(k,6) = L_O(k);
O(k,7) = D_O(k);
O(k,8) = LL_O(k);

S(k,2) = T_S(k);
S(k,3) = B_S(k);
S(k,4) = P_S(k);
S(k,5) = L_S(k);
S(k,6) = D_S(k);
S(k,7) = LL_S(k);

S(k,1) = (k-1);
I(k,1) = (k-1);
E(k,1) = (k-1);
O(k,1) = (k-1);
EX(k,1) = (k-1);
q(k,1) = (k-1);

%%% Shallow and Deep Aquifer

% calculating percolation of shallow aquifer
F(k) = (x(6) - MSA(k-1))./(x(6)).*x(10);
if F(k) < 0;
    F(k) = 0;
end

% mass balance shallow aquifer
MSA(k) = MSA(k-1) - I(k,2)+ EX(k,2) + E(k,8) + F(k);
if MSA(k) > (x(11)*x(12));
    EX(k,2) = 0;
    F(k) = 0;
    MSA(k) = MSA(k-1) - I(k,2)+ EX(k,2) + E(k,8) + F(k);
end

% mass balance deep aquifer
MDA(k) = MDA(k-1) - F(k) + PU(k);

%%% Discharge

% 1 2    3   4    5      6
% k tot mes quick slow quick2

% Overflow quick
if k>x(3)
    if x(3)-round(x(3)) ~= 0
```

```matlab
            q(k,6) = -alphaq.*q(k-1,6) + betaq.*((ceil(x(3))-x(3))...
                .*O(k-floor(x(3)),2)+(x(3)-floor(x(3))).*...
                O(k-ceil(x(3)),2)).*(x(5)./100);
        else
            q(k,6) = -alphaq.*q(k-1,6) + betaq.*O(k-x(3),2).*(x(5)./100);
        end
        q(k,4) = -alphaq.*q(k-1,4) + betaq.*q(k,6);
    end

    % Overflow slow
    if k>x(4)
        if x(4)-round(x(4))~= 0
            q(k,5) = -alphas.*q(k-1,5) + betas.*((ceil(x(4))-x(4))...
                .*O(k-floor(x(4)),2)+(x(4)-floor(x(4))).*...
                O(k-ceil(x(4)),2)).*((100-x(5))./100);
        else
            q(k,5) = -alphas.*q(k-1,5) + betas.*O(k-x(4),2).*((100-x(5))./100);
        end
    end

    % Total
    q(k,2) = (q(k,4)+q(k,5))./1000./3600./24.*A; % mm/dag -> m3/s
    q(k,3) = Q(k-1);
end

%%% Calculating NS and RVE
w = t1:t2;
% Starting empty NS and RVE calculation
NS_num = zeros(length(w),1);
NS_denom = zeros(length(w),1);
RVE_num = zeros(length(w),1);
RVE_denom = zeros(length(w),1);

for w = t1:t2
    if q(w,3) >= 0
        NS_num(w) = (q(w,2) - q(w,3)).^2;
        NS_denom(w) = (q(w,3) - mean(q(:,3))).^2;
        RVE_num(w) = q(w,2) - q(w,3);
        RVE_denom(w) = q(w,3);
    end
end

NS = 1-(sum(NS_num)./sum(NS_denom));
RVE = 100.*(sum(RVE_num)./sum(RVE_denom));

NS,RVE   %#ok<NOPTS>

%%%%%%% Producing matrices for graphs %%%%%%%%
% X:        x-values
% Y1:       q modelled
% Y2:       q measured
% Y3:       Storage
% Y4:       Exfiltration

% Y6:       Evapotranspiration Surface
% Y7:       Evapotranspiration shallow aquifer
% Y8:       Total Evapotranspiration

% Producing subtotals for graphs on:
% Mes:      q measured
% Mod:      q modelled
% Rain:     Rainfall
% E:        Evapotranspiration
% EX:       Exfiltration

X = (0:k-1)';                       % x-values for graphs
Y1 = q(1:k,2)./A.*86400000;         % q modelled (m3/s --> mm/dag)
Y1_S = q(:,5);                      % component slow overflow (mm/dag)
Y1_Q = q(:,4);                      % component quick overflow (mm/dag)
Y2 = q(1:k,3)./A*86400000;          % q measured
Y2_cum = Data1(:,13)./A.*86400000;  % q measured for cumulative graph (with 0 values in stead of negative
values)

RR_coefficient = sum(Y1(246:2556))./sum(P(246:2556)) %#ok<NOPTS>

% starting empty for subtotals calculation
Subtotal_Mes = zeros(k,1);
Subtotal_Mod = zeros(k,1);
Subtotal_Rain = zeros(k,1);
Subtotal_E = zeros(k,1);

for i = t1:t2+1
    Subtotal_Mes(i) = Subtotal_Mes(i-1) + Y2_cum(i);
    Subtotal_Mod(i) = Subtotal_Mod(i-1) + Y1(i);
    Subtotal_Rain(i) = Subtotal_Rain(i-1) + P(i);
    Subtotal_E(i) = Subtotal_E(i-1) + E(i,2) + E(i,8);
end
```

```matlab
%%% Extra graph calculations
Y3_B = S(:,3);            % Baidh storage
Y3_P = S(:,4);            % Pond storage
Y3_L = S(:,5);            % Lowland storage
Y3_D = S(:,6);            % Dam storage
Y3_LL = S(:,7);           % LLowland storage
Y4_L = EX(:,5);           % Lowland exfiltration
Y4_D = EX(:,6);           % Dams Exfiltration
Y4_LL = EX(:,4);          % LLowlands Exfiltration
Y4_P = EX(:,7);           % Ponds Exfiltration
Y6 = E(:,2);              % Evapotranspiration Surface
Y7 = E(:,8);              % Evapotranspiration Shallow Aquifier
Y8 = E(:,8) + E(:,2);     % total evapotranspiration

% Suggestions on figures
% fig.  x: modelled and measured runoff graphs
% fig. 1x: subtotals graphs
% fig. 2x: storage graphs
% fig. 3x: exfiltration and sub surface flow graphs

% option to shut off the graphs, increases processing time
% {
figure(1), clf
plot(X,Y2(1:size(X)),X,P(1:size(X)),X,Y1(1:size(X)))
title('Modelled output compared with observed measurements and rainfall')
xlim([2436 2556])
ylim([0 25])
legend('Observed flow (mm)','Rainfall (mm)','Modelled flow (mm)')
xlabel('1 February 1996 - 31 May 1996')
ylabel('Quantity (mm)')

figure(2), clf
plot(X,Y1_S(1:size(X)),X,Y1_Q(1:size(X)),X,Y1(1:size(X)),':')
xlim([t1 t2])
title('Modelled runoff splitted in quick and slow component')
legend('Slow modelled runoff component','Quick modelled runoff component','Total modelled runoff')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('stream flow (mm)')

figure(11), clf
plot(X,Subtotal_Mod,X,Subtotal_Mes)
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('stream flow (mm)')
title('Cumulative values modelled and observed stream flow')
legend('Modelled streamflow','Observed streamflow','location','SouthEast')

figure(12), clf
plot(X,Subtotal_Rain,X,Subtotal_E(:,1))
xlim([t1 t2])
title('Cumulative graphs rainfall and evapotranspiration')
legend('rainfall','Evapotranspiration','location','SouthEast')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('mm')

figure(21), clf
plot(X,Y3_B(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')
title('Baidh storage')

figure(22), clf
plot(X,Y3_P(1:size(X)))
xlim([t1 t2])
title('Pond storage')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')

figure(23), clf
plot(X,Y3_L(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')
title('Lowland storage')

figure(24), clf
plot(X,Y3_D(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')
title('Dam storage')

figure(25), clf
plot(X,Y3_LL(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')
```

```matlab
title('Lower Lowland storage')
legend('LLowland storage')

figure(26), clf
plot(X,MSA(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('deficit in mm')
title('Deficit shallow aquifer')
legend('MSA')

figure(31), clf
plot(X,F(1:size(X)))
xlim([t1 t2])
title('Percolation')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(32), clf
plot(X,Y4_LL(1:size(X)))
xlim([t1 t2])
title('Exfiltration Lower Lowlands')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(33), clf
plot(X,Y4_L(1:size(X)))
xlim([t1 t2])
title('Exfiltration Lowlands')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(34), clf
plot(X,Y4_D(1:size(X)))
xlim([t1 t2])
title('Exfiltration Dams')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(35), clf
plot(X,Y6(1:size(X)),X,pE(1:size(X)))
xlim([t1 t2])
title('Evaporation surface')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')
legend('Evaporation')

figure(36), clf
plot(X,Y7(1:size(X)),X,pE(1:size(X)))
xlim([t1 t2])
title('Evapotranspiration from shallow Aquifier')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')
legend('Evapotranspiration')

figure(37), clf
plot(X,Y8(1:size(X)),X,pE(1:size(X)))
xlim([t1 t2])
title('Total evapotranspiration')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')
legend('Evapotranspiration')
%}
```
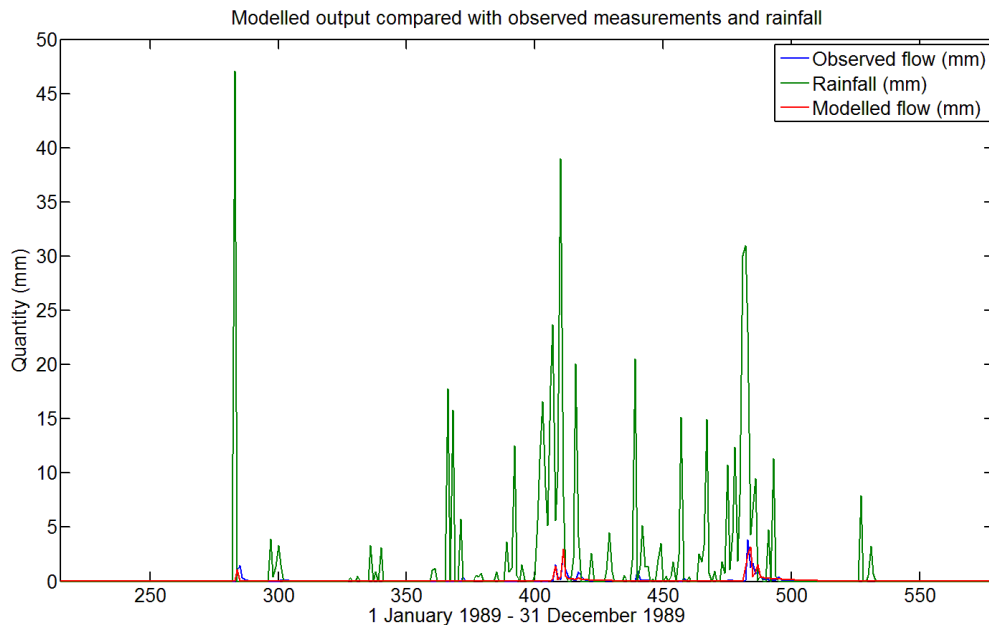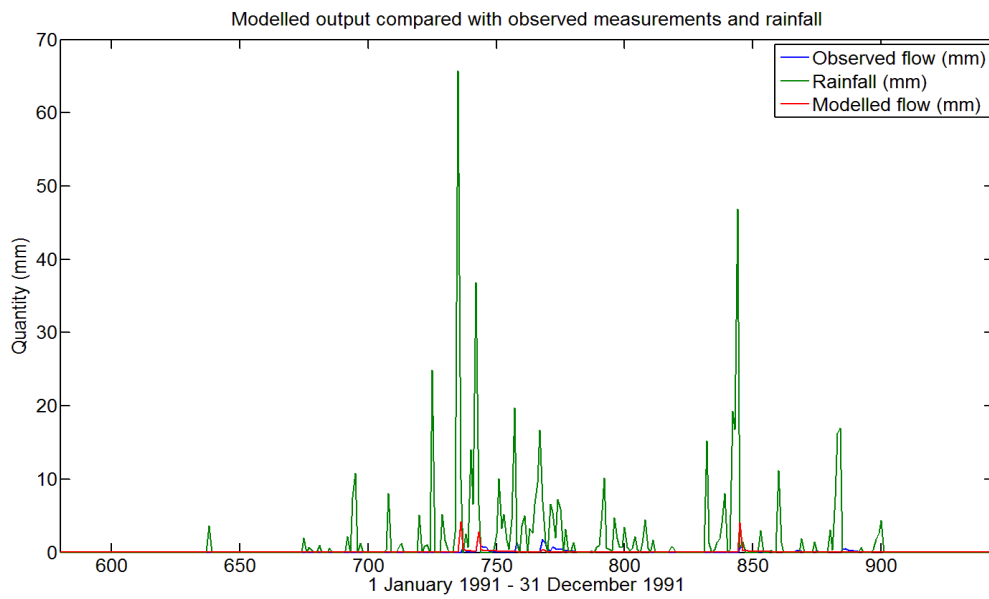
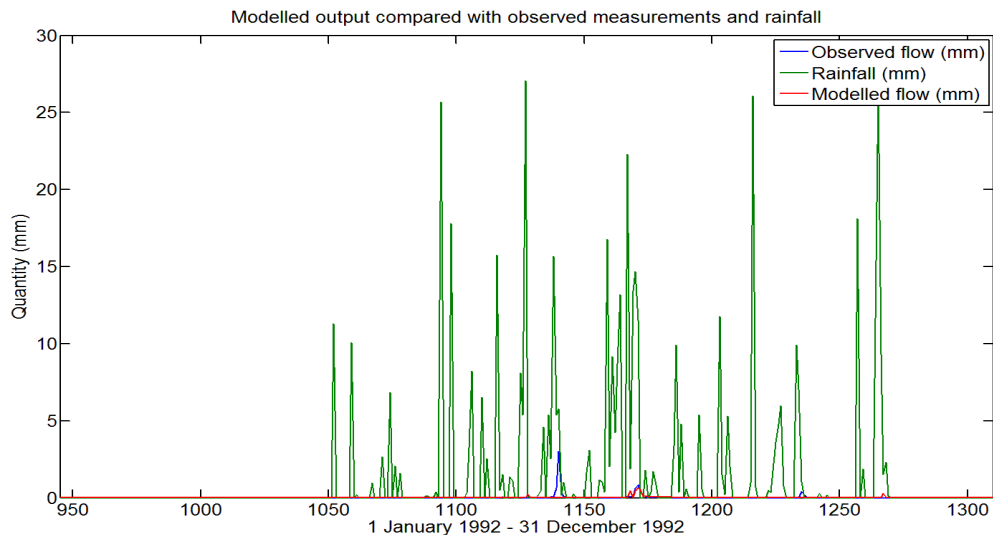## C. Yearly modelled results for gauged catchment Andhra Pradesh
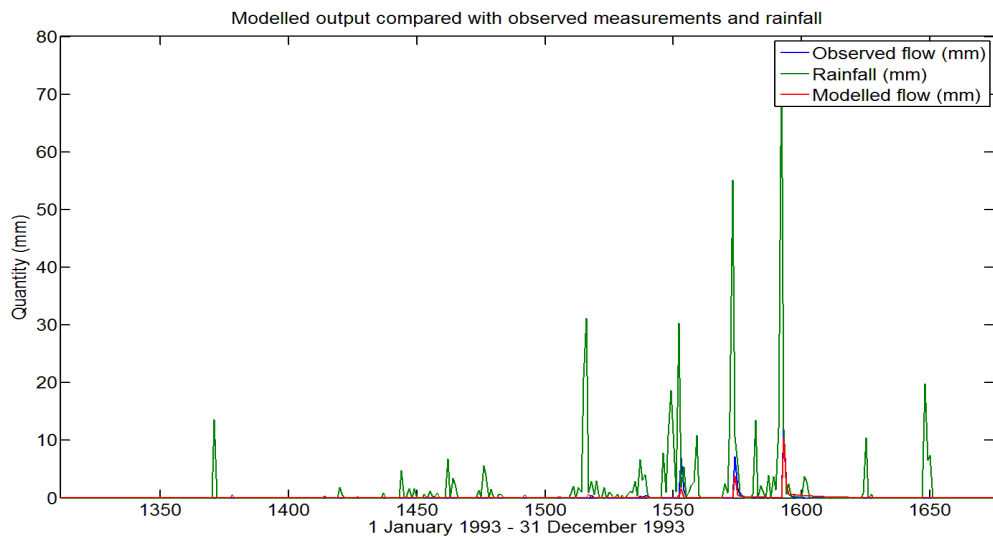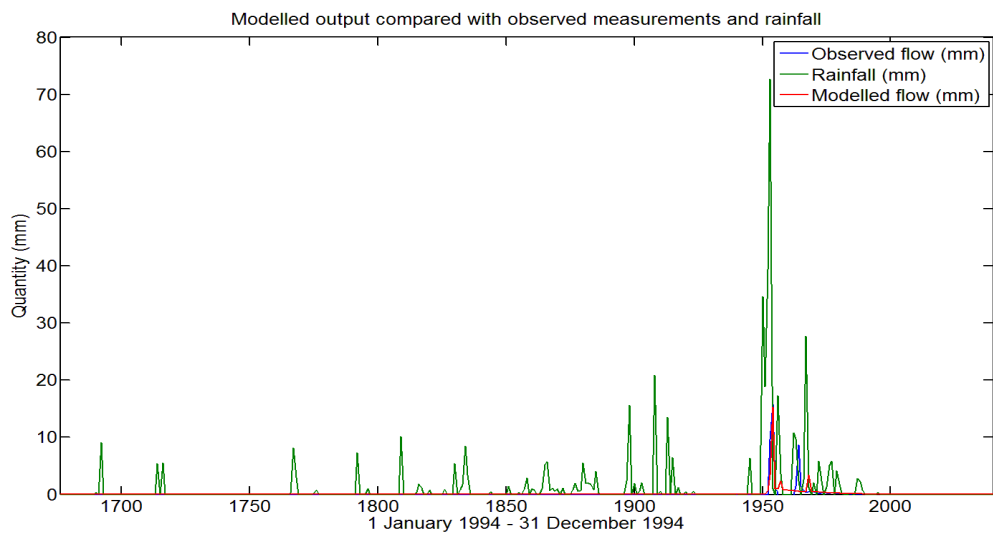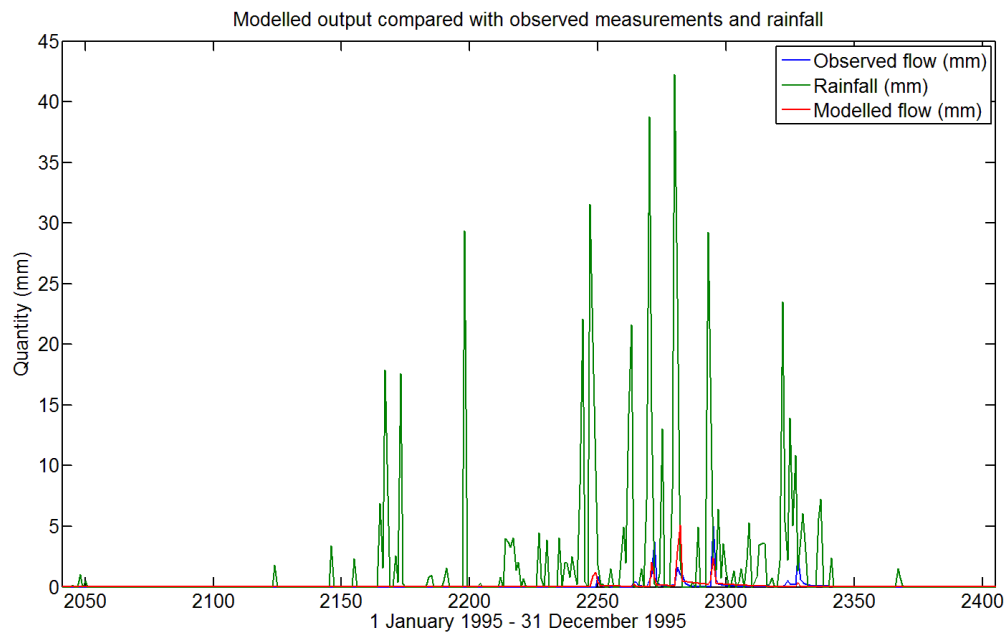
Year 1989:

Modelled output compared with observed measurements and rainfall

Year 1991:

Modelled output compared with observed measurements and rainfall

Year 1992:

Year 1993:

Year 1994:

Year 1995:

Modelled output compared with observed measurements and rainfall

1 January 1995 - 31 December 1995

Year 1996:

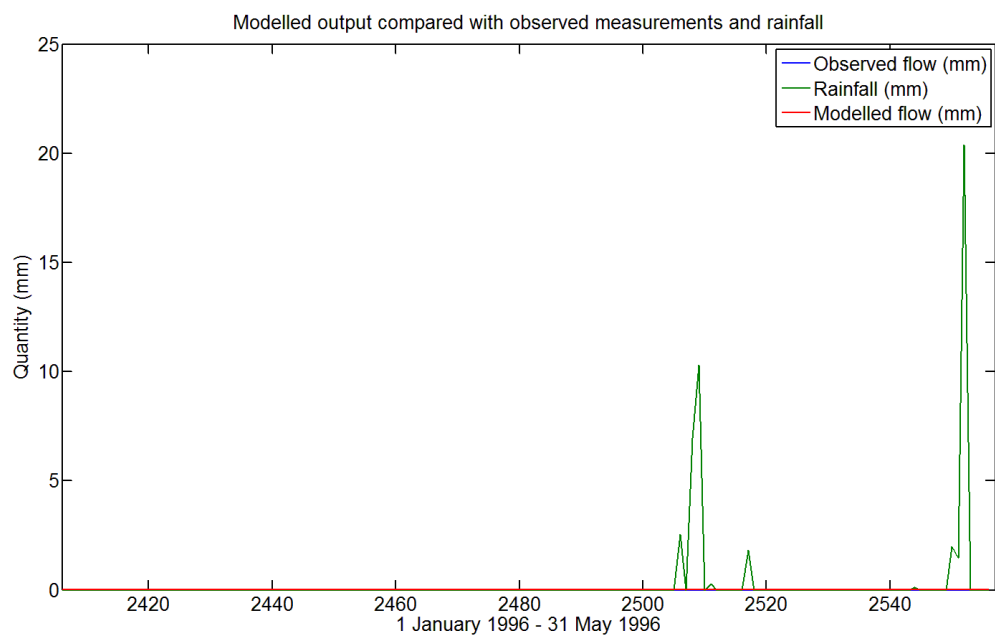Modelled output compared with observed measurements and rainfall

1 January 1996 - 31 May 1996

D. Matlab code ungauged Gooty study site

```matlab
% Model Gooty ungauged study site by Rinse Wilmink
clear all
close all
clc

%---- Input Data ----------------------------------------------
load input_Gooty.csv
Data1 = input_Gooty;

%Data 1: rainfall,temp,day of month
%         (mm)     ,C   , nr

x  = [0.0491 12.6 0   0  54.5 1725  1725 910  4   120  10000  0.4];
%      x1      x2  x3  x4 x5    x6    x7   x8  x9   x10  x11    x12
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% x1 Tau quick
% x2 Tau slow
% x3 Delay quick
% x4 Delay slow
% x5 Proportion quick (100%)
% x6 Threshold MSA F
% x7 Threshold MSA EX (g in report)
% x8 Threshold MSA eP SA (h in report)
% x9 Exfiltration parameter (mm/dag) (G in report)
%x10 Percolation parameter (mm/dag)  (F in report)
%x11 depth shallow aquifer (mm)
%x12 porosity saturated shallow aquifer
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% t1 and t2 for period selection (t1 to t2) graphs
% Whole data set: set to: k = 2:size(Data,1)
t1 = 215;     % originally t1 = 2 (start dataset), t1 215 = 1 January 1989
t2 = 2556;    % originally t2 = 2557-1 ( end dataset)
n = 1;        % startpoint FOR-loop and initial values MSA, MDA

% input: rainfall, pet
Zeros = zeros(size(Data1,1),1); %starting empty

temp = Data1(:,2);
P = Data1(:,1);
pE = temp.*0.1446;

timestep = 24;   % timestep in days, 24 hour
A = 205.*10^6;   % Area A is 205 sqkm

% Surface Module Parameters
%%% Tarh (T)
T_Smax = 3;      % 3 originally;
T_Ksat = 30;
T_Ac = 0.25;
%%% Baidh (B)
B_Smax = 11;     % 50 originally;
B_Ksat = 0.5;
B_Ac = 0.39;
%%% Ponds (P)
P_Smax = 3000;   % 3000 originally;
P_Ksat = 0.4;
P_Ac = 0.01;
%%% Lowland (L)
L_Smax = 126;     % 100 originally;
L_Ksat = 0.05;
L_Ac = 0.35;

MSA = Zeros;      % Moisture deficit shallow aquifer
MSA(n) = 1775;    % Start value
MDA = Zeros;      % Moisture deficit deep aquifer
MDA(n) = 2000;    % Start value

%%% Discharge q parameters
tauq = (x(1));       %(day)
taus = (x(2));       %(day)

alphaq = -exp(-1./tauq);
betaq = 1+alphaq;
alphas = -exp(-1./taus);
betas = 1+alphas;

T_Ksat = T_Ksat .* timestep;
B_Ksat = B_Ksat .* timestep;
P_Ksat = P_Ksat .* timestep;
L_Ksat = L_Ksat .* timestep;

T_S = Zeros; %starting empty
T_O = Zeros;
T_E = Zeros;
```

```matlab
T_I = Zeros;
B_S = Zeros; %starting empty
B_O = Zeros;
B_E = Zeros;
B_I = Zeros;
P_S = Zeros; %starting empty
P_O = Zeros;
P_E = Zeros;
P_I = Zeros;
L_S = Zeros; %starting empty
L_O = Zeros;
L_E = Zeros;
L_I = Zeros;
F = Zeros;        % Peroclation from shallow aquifer
PU = Zeros;       % Pumping

% Area fractions surface evapotranspiration
T_fs = Zeros;
B_fs = Zeros;
L_fs = Zeros;

I = zeros(length(Zeros),6);
E = zeros(length(Zeros),8);
O = zeros(length(Zeros),6);
S = zeros(length(Zeros),5);
EX = zeros(length(Zeros),6);
q = zeros(length(Zeros),6);

L_AcE = 0;
B_AcE = 0;
T_AcE = 0;

%%% Calculation
for k = (n+1):size(Data1(:,1)) % whole dataset

    %%% Exfiltration
    if MSA(k-1) < x(7);
        EX(k,2) = x(9).*((x(7) - MSA(k-1))./x(7));
    else
        EX(k,2) = 0;
    end
    EX(k,3) = min(EX(k,2),L_Ksat.*L_Ac);
    if EX(k,2)-EX(k,3) > 0
        EX(k,4) = min(EX(k,2)-EX(k,3),B_Ksat.*B_Ac);
        if EX(k,2)-(EX(k,3)+EX(k,4)) > 0
            EX(k,5) = min(EX(k,2)-(EX(k,3)+EX(k,6)+...
                EX(k,7)),T_Ksat.*T_Ac);
            if EX(k,2)-(EX(k,3)+ EX(k,4)+EX(k,5)) > 0
                EX(k,6) = EX(k,2)-(EX(k,3)+EX(k,4)+...
                    EX(k,5));
            end
        end

    end
    L_AcE = EX(k,3)./(L_Ksat.*L_Ac);
    B_AcE = EX(k,4)./(B_Ksat.*B_Ac);
    T_AcE = EX(k,5)./(T_Ksat.*T_Ac);

    %%% Calculating evaportanspiration shallow aquifer
    E(k,8) = (pE(k-1)).*min(1,exp(2.*(1-MSA(k-1)/x(8))));

    %%% Pumping from deep aquifer
    PU_days = [0:45,280:365,610:695,1005:1090,1380:1445,1745:1830,2135:220]; % pumping day numbers
    if ismember(k,PU_days) == 1;
        PU(k) = 0.262;
    else
        PU(k) = 0.0;
    end

    %%% Tarh (T)
    T_S(k) = T_S(k-1) + P(k-1);
    T_I(k) = min(T_S(k),T_Ksat.*(1-T_AcE));
    T_S(k) = T_S(k) - T_I(k) + EX(k,5)./T_Ac;
    T_O(k) = max(T_S(k)-T_Smax,0);
    T_S(k) = T_S(k) - T_O(k);

    T_u = 0.8;
    T_fs(k) = exp(-(log(T_S(k)./T_Smax)./log(T_u)).^2);

    T_E(k) = min(T_S(k),T_fs(k).*(pE(k-1)-E(k,8)));
    T_S(k) = T_S(k) - T_E(k);

    %%% Baidh (B)
    B_S(k) = B_S(k-1) + P(k-1) + T_O(k).*T_Ac./B_Ac + PU(k)./B_Ac;
    B_I(k) = min(B_S(k),B_Ksat.*(1-B_AcE));
    B_S(k) = B_S(k) - B_I(k) + EX(k,4)./B_Ac;
    B_O(k) = max(B_S(k)-B_Smax,0);
    B_S(k) = B_S(k) - B_O(k);
```

```matlab
B_u = 0.6;
B_fs(k) = exp(-(log(B_S(k)./B_Smax)./log(B_u)).^2);


B_E(k) = min(B_S(k),B_fs(k).*(pE(k-1)-E(k,8)));
B_S(k) = B_S(k) - B_E(k);

%%% Ponds (P)
P_S(k) = P_S(k-1) + P(k-1) + B_O(k).*B_Ac./P_Ac;
P_I(k) = min(P_S(k),P_Ksat);
P_S(k) = P_S(k) - P_I(k);
P_O(k) = max(P_S(k)-P_Smax,0);
P_S(k) = P_S(k) - P_O(k);
P_E(k) = min(P_S(k),(pE(k-1)-E(k,8)));
P_S(k) = P_S(k) - P_E(k);

%%% Lowland (L)
L_S(k) = L_S(k-1) + P(k-1) + P_O(k).*P_Ac./L_Ac;
L_I(k) = min(L_S(k),L_Ksat.*(1-L_AcE));
L_S(k) = L_S(k) - L_I(k) + EX(k,3)./L_Ac;
L_O(k) = max(L_S(k)-L_Smax,0);
L_S(k) = L_S(k) - L_O(k);

L_u = 0.5;
L_fs(k) = exp(-(log(L_S(k)./L_Smax)./log(L_u)).^2);

L_E(k) = min(L_S(k),L_fs(k).*(pE(k-1)-E(k,8)));
L_S(k) = L_S(k) - L_E(k);

%%% Surface module output Infiltration, Evaporation, Overflow, Storage
I(k,2) = T_I(k).*T_Ac + B_I(k).*B_Ac + P_I(k).*P_Ac + L_I(k).*L_Ac;
I(k,3) = T_I(k);
I(k,4) = B_I(k);
I(k,5) = P_I(k);
I(k,6) = L_I(k);

E(k,2) = T_E(k).*T_Ac + B_E(k).*B_Ac + P_E(k).*P_Ac + L_E(k).*L_Ac;
E(k,3) = T_E(k);
E(k,4) = B_E(k);
E(k,5) = P_E(k);
E(k,6) = L_E(k);
%E(k,8) = shallow aquifer evapotranspiration

O(k,2) = L_O(k).*L_Ac;
O(k,3) = T_O(k);
O(k,4) = B_O(k);
O(k,5) = P_O(k);
O(k,6) = L_O(k);

S(k,2) = T_S(k);
S(k,3) = B_S(k);
S(k,4) = P_S(k);
S(k,5) = L_S(k);

S(k,1) = (k-1);
I(k,1) = (k-1);
E(k,1) = (k-1);
O(k,1) = (k-1);
EX(k,1) = (k-1);
q(k,1) = (k-1);

%%% Shallow and Deep Aquifer
% calculating percolation of shallow aquifer
F(k) = (x(6) - MSA(k-1))./(x(6)).*x(10);
if F(k) < 0;
    F(k) = 0;
end

% mass balance shallow aquifer
MSA(k) = MSA(k-1) - I(k,2)+ EX(k,2) + E(k,8) + F(k);
if MSA(k) > (x(11)*x(12));
    EX(k,2) = 0;
    F(k) = 0;
    MSA(k) = MSA(k-1) - I(k,2)+ EX(k,2) + E(k,8) + F(k);
end

% mass balance deep aquifer
MDA(k) = MDA(k-1) - F(k) + PU(k);

%%% Discharge
% 1 2   3   4     5       6
% k tot mes quick slow quick2

% Overflow quick
if k>x(3)
    if x(3)-round(x(3)) ~= 0
        q(k,6) = -alphaq.*q(k-1,6) + betaq.*((ceil(x(3))-x(3))...
            .*O(k-floor(x(3)),2)+(x(3)-floor(x(3))).*...
```

```matlab
                O(k-ceil(x(3)),2)).*(x(5)./100);
        else
            q(k,6) = -alphaq.*q(k-1,6) + betaq.*O(k-x(3),2).*(x(5)./100);
        end
        q(k,4) = -alphaq.*q(k-1,4) + betaq.*q(k,6);
    end

    % Overflow slow
    if k>x(4)
        if x(4)-round(x(4))~= 0
            q(k,5) = -alphas.*q(k-1,5) + betas.*((ceil(x(4))-x(4))...
                .*O(k-floor(x(4)),2)+(x(4)-floor(x(4))).*...
                O(k-ceil(x(4)),2)).*((100-x(5))./100);
        else
            q(k,5) = -alphas.*q(k-1,5) + betas.*O(k-x(4),2).*((100-x(5))./100);
        end
    end

    % Total
    q(k,2) = (q(k,4)+q(k,5))./1000./3600./24.*A; % mm/dag -> m3/s
 end

%%%%%%%% Producing matrices for graphs %%%%%%%%
% X:        x-values
% Y1:       q modelled
% Y3:       Storage of Ponds (Y3_P), Baidh (Y3_B) and Lowland (Y3_L)
% Y4:       Exfiltration into Lowland (Y4_L) and Baidh (Y4_B)
% Y6:       Evapotranspiration Surface
% Y7:       Evapotranspiration shallow aquifer
% Y8:       Total Evapotranspiration

% Producing subtotals for graphs on:
% Mes:      q measured
% Mod:      q modelled
% Rain:     Rainfall
% E:        Evapotranspiration
% EX:       Exfiltration

X = (0:k-1)';                   % x-values for graphs
Y1 = q(1:k,2)./A.*86400000;     % q modelled (m3/s --> mm/dag)
Y1_S = q(:,5);                  % component slow overflow (mm/dag)
Y1_Q = q(:,4);                  % component quick overflow (mm/dag)

RR_coefficient = sum(Y1(246:2556))./sum(P(246:2556)) %#ok<NOPTS>

Subtotal_Mod = zeros(k,1);
Subtotal_Rain = zeros(k,1);
Subtotal_E = zeros(k,1);

for i = t1:t2+1
    Subtotal_Mod(i) = Subtotal_Mod(i-1) + Y1(i);
    Subtotal_Rain(i) = Subtotal_Rain(i-1) + P(i);
    Subtotal_E(i) = Subtotal_E(i-1) + E(i,2) + E(i,8);
end

%%% Extra graph calculations
Y3_B = S(:,3);              % Baidh storage
Y3_P = S(:,4);              % Pond storage
Y3_L = S(:,5);              % Lowland storage
Y4_L = EX(:,3);            % Lowland exfiltration
Y4_B = EX(:,4);            % Baidh exfiltration
Y6 = E(:,2);                % Evapotranspiration Surface
Y7 = E(:,8);                % Evapotranspiration Shallow Aquifer
Y8 = E(:,8) + E(:,2);      % total evapotranspiration

% Suggestions on figures
% fig.  x: modelled and measured runoff graphs
% fig. 1x: subtotals graphs
% fig. 2x: storage graphs
% fig. 3x: exfiltration and sub surface flow graphs

figure(1), clf
plot(X,P(1:size(X)),X,Y1(1:size(X)))
title('Modelled output')
xlim([2437 2556])
ylim([0 10])
legend('Rainfall (mm)','Modelled flow (mm)')
xlabel('1 January 1989 - 31 May 1996')
ylabel('Quantity (mm)')

figure(2), clf
plot(X,Y1_S(1:size(X)),X,Y1_Q(1:size(X)),X,Y1(1:size(X)),':')
xlim([t1 t2])
title('modelled streamflow splitted out in quick and slow component')
legend('slow component','quick component','Total')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')

figure(11), clf
```

```
plot(X,Subtotal_Mod)
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('stream flow (mm)')
title('Cumulative values modelled stream flow')
legend('Modelled streamflow','location','SouthEast')

figure(12), clf
plot(X,Subtotal_Rain,X,Subtotal_E(:,1))
xlim([t1 t2])
title('Cumulative values rainfall and evapotranspiration')
legend('rainfall','Evapotranspiration','location','SouthEast')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('mm')

figure(21), clf
plot(X,Y3_B(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')
title('Baidh storage')

figure(22), clf
plot(X,Y3_P(1:size(X)))
xlim([t1 t2])
title('Pond storage')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')

figure(23), clf
plot(X,Y3_L(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('storage in mm')
title('Lowland storage')

figure(26), clf
plot(X,MSA(1:size(X)))
xlim([t1 t2])
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('deficit in mm')
title('Deficit shallow aquifer')
legend('MSA')

figure(31), clf
plot(X,F(1:size(X)))
xlim([t1 t2])
title('Percolation')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(32), clf
plot(X,Y4_L(1:size(X)))
xlim([t1 t2])
title('Exfiltration Lowlands')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(33), clf
plot(X,Y4_B(1:size(X)))
xlim([t1 t2])
title('Exfiltration Baidh')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')

figure(34), clf
plot(X,Y6(1:size(X)),X,pE(1:size(X)))
xlim([t1 t2])
title('Evaporation surface')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')
legend('Evaporation')

figure(35), clf
plot(X,Y7(1:size(X)),X,pE(1:size(X)))
xlim([t1 t2])
title('Evapotranspiration from shallow Aquifier')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')
legend('Evapotranspiration')

figure(36), clf
plot(X,Y8(1:size(X)),X,pE(1:size(X)))
xlim([t1 t2])
title('Total evapotranspiration')
xlabel('Dataset daynumbers (1 January 1989 - 31 December 1989 and 1 January 1991 - 31 May 1996)')
ylabel('flow in mm/timestep')
legend('Evapotranspiration')
```
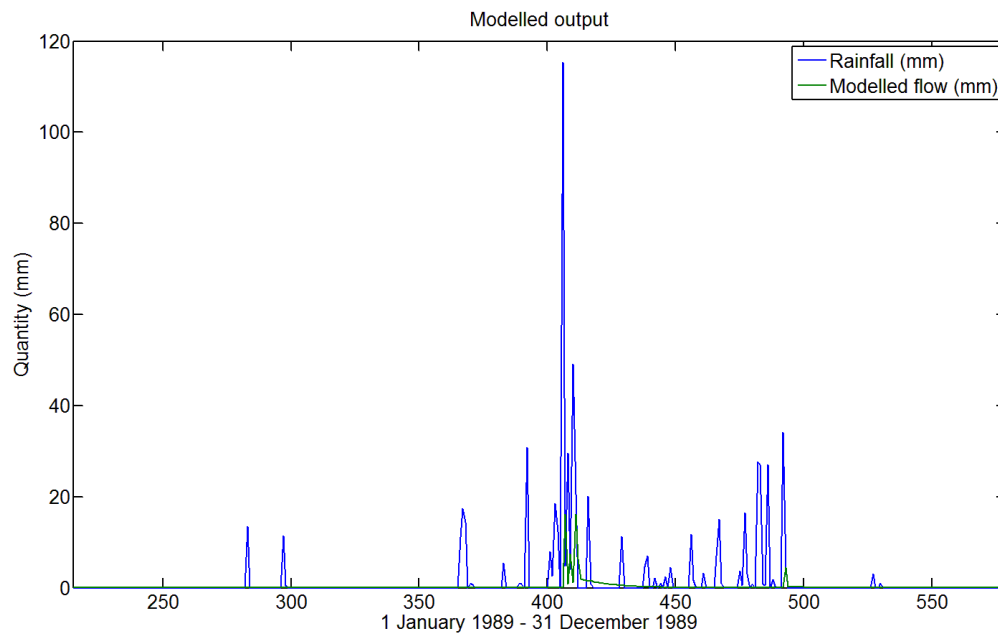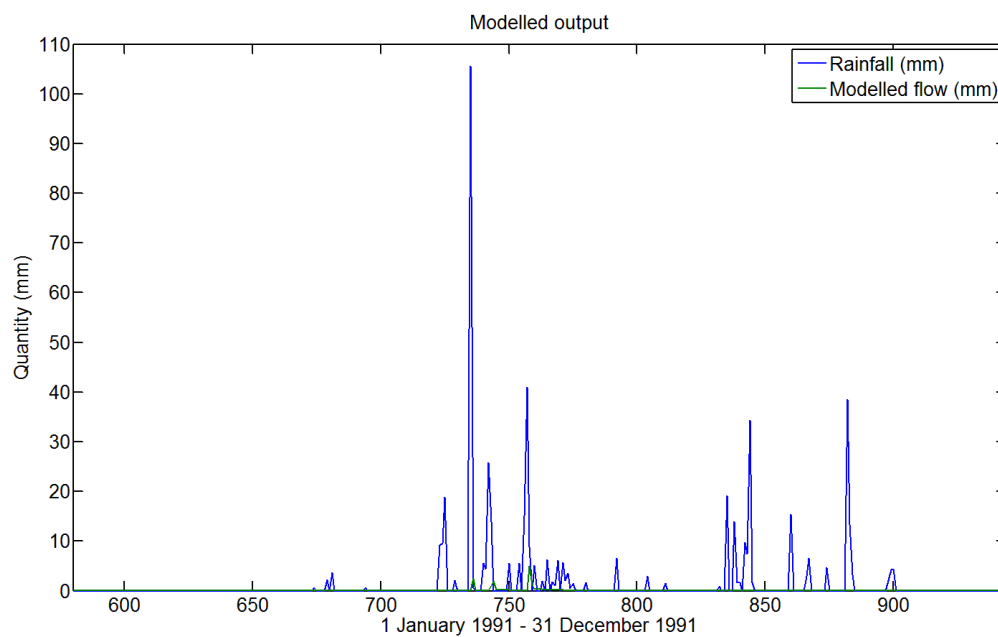
E. Yearly modelled results for ungauged Gooty study site
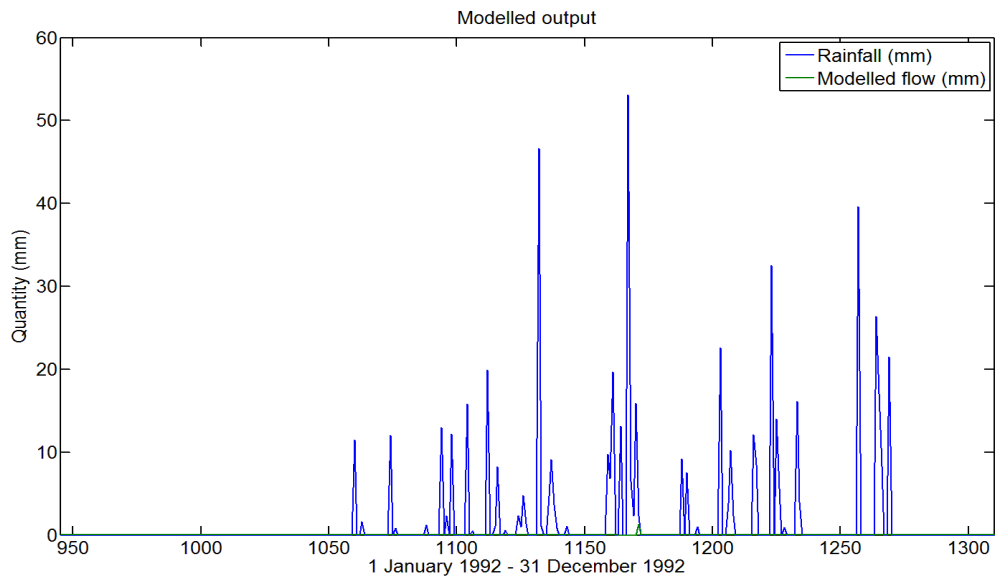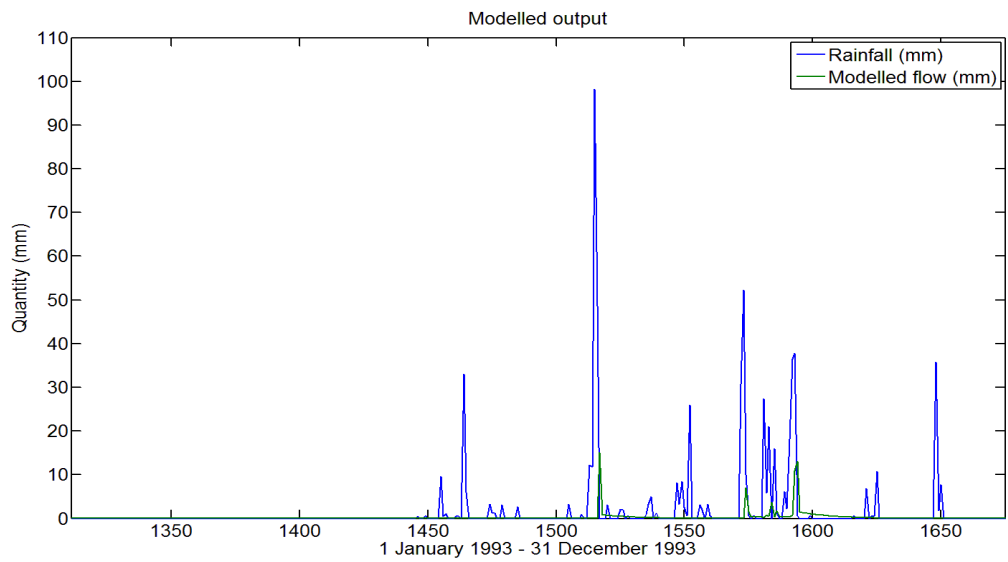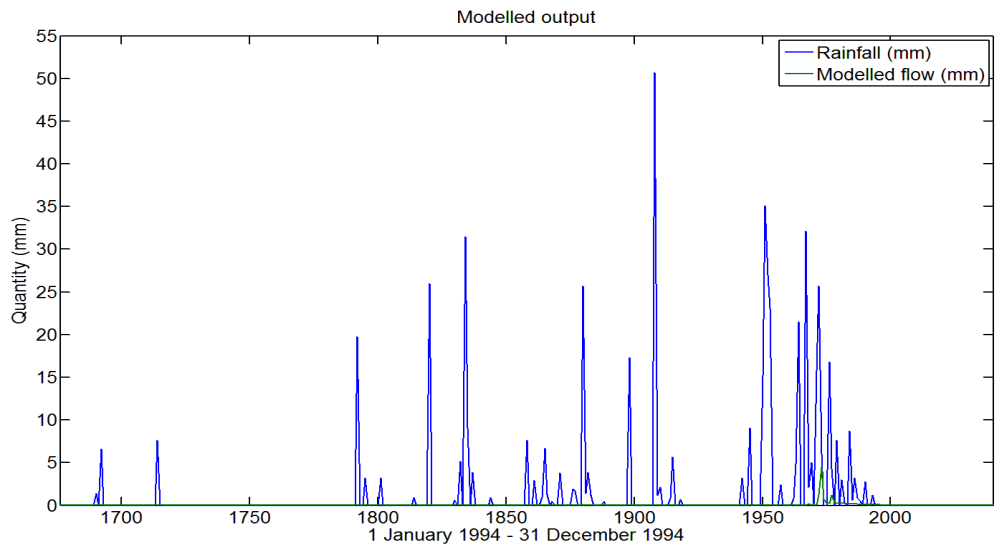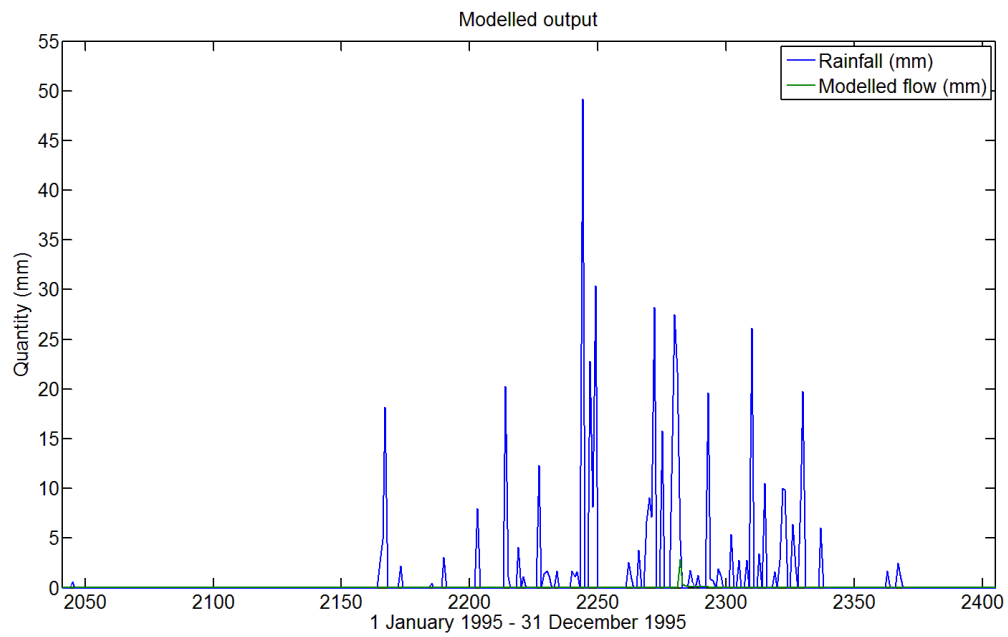
Year 1989:

Year 1991:

Year 1992:

Year 1993:

Year 1994:

Year 1995:

Year 1996: