Physiographical baseline survey for the Upper Tana catchment: erosion and sediment yield assessment

May 2011

Author
J.E. Hunink, P. Droogers

Client
Water Resources Management Authority (WRMA), Kenya

Report FutureWater: 112

FutureWater
Costerweg 1V
6702 AA Wageningen
The Netherlands

+31 (0)317 460050

info@futurewater.nl

www.futurewater.nl

Preface

This consultancy undertook an assessment of the soil erosion and sediment loads in the Upper Tana catchment areas and estimated the impact of sediment deposition on the reservoir capacities of the principal dams. A physiographical baseline survey was carried out, including a detailed monitoring campaign of flow and sediment loads throughout the basin, bathymetric surveys of the reservoirs and soil erosion modeling to assess the current status of the basin. This report summarizes the erosion and sediment yield modeling assessment carried out by FutureWater. It is part of the full project report “Consultancy Services for the Physiographical Survey in the Upper Tana Catchment, Loan No4772-KE”, by Z&A P. Antonaropoulos and Associates S.A. and G.Karavokyris & Partners Consulting Engineers S.A., Athens, Greece, for the Water Resources Management Authority (WRMA), Kenya, and with funds of the World Bank.

Table of contents

Tables

Figures

1 Introduction

The Physiographical Baseline Survey of the Upper Tana catchment, implemented by the Water Resources Management Authority (WRMA) of Kenya, included an extensive monitoring campaign and bathymetric surveys, which gave an integral picture of the current and historic erosion and the sediment budgets in the basin. To obtain more information on the spatial and temporal variability of the sediment dynamics in the basin, distributed erosion and hydrological modelling was carried out to complement the measurements and to study future scenarios. Modelling allows quantifying the erosion processes at non-gauged areas and during periods when measurements were absent or poor. Also different future management scenarios can be studied and compared to support following-up decision on possible interventions. This report summarizes the results from the modelling assessment. In the appendices, detailed model output (model performance, maps, tables) can be found.

The model was calibrated using a for this basin unprecedented amount of data covering 30 years from different locations. The bathymetric surveys performed within this project yielded precise data on the siltation rate which was reproduced by simulations of the model that was run from when they started to be operational and started trapping sediments until 2011. The model was validated with a dataset that is most favourable for its location, scale and consistency, being the Masinga total inflow, calculated from mass balance (levels and releases).

The last sections of this report show the results of the scenario analysis, and of the targeting of the WRUAs for implementation of these scenarios. Also, different adoption levels were assessed to support decisions to be made on the scale of implementation in the following-up phase.

2 Model setup

2.1 Model selection

The hydrological and erosion assessment tool selected for this project was required to allow the simulation of the erosion processes for all combinations of soil types and land uses available in the basin, with a relatively high level of spatial detail. The hydrological model that was considered most suitable for this project is the Soil Water Assessment Tool (SWAT), being used by a large user group all over the world for similar assessments (e.g. Tripathi et al, 2003; Parajuli et al, 2008).

The main strength for this particular project is that the tool uses a physical based rainfall-runoff scheme, instead of a data-based statistic or conceptual scheme, used by many other models used for basin-scale assessments. This guarantees more reliable scenario simulations and better performance in poorly gauged catchments, which is essential for a study on this scale. Besides, the model is primarily focused on the interaction between land management versus water- and erosion processes. This is useful as the major part of gross erosion is related to practicing certain land management. Several previous hydrological studies were carried out in Tana Basin of which a few ones used SWAT as the main assessment tool.

Shortly, strong aspects of the SWAT model that make the tool suitable for this project can be summarized as follows:

- Physically based rather than parametric based rainfall-runoff scheme to ensure more reliable scenario simulations.
- Developed with a focus on water-erosion-land management processes.
- Software is public domain, including source code is freely available.
- Integrated in mainstream GIS tools with a user-friendly interface, making it accessible to technical staff of local authorities.
- It counts with a large user-group world wide, providing support.
- Excellent documentation is provided, including training materials.
- The consortium carrying out this contract has extensive experiences in application as well as training of this particular model.
- Previous modeling experience in this basin through the Green Water Credits project using the same tool, focusing on Blue and Green water flows but also on erosions and sediment transport.

Being a physically based model, the data requirements are relatively high. This means that it is necessary to dedicate substantial efforts to data collection and data preparation. As this project includes many data collection activities, SWAT allows optimal usage of these data. Not only the data measured within the current project will be used, but also several other spatial datasets on land use, soil types and characteristics, elevation, rainfall distribution, etc, which have been obtained from local and public domain sources.

Another difficulty to deal with using SWAT is the balance between model run time and the amount of spatial detail included in the model configuration. SWAT is a model designed for basin-scale assessments, but it has the capability to include almost an unlimited amount of spatial detail. This flexibility is of course particularly useful for this project as erosion is a spatially highly heterogeneous process of which the determining parameters can change within very short distances. However, due to computational constraints, the amount of calculation

nodes have to be limited to a reasonable number, For this project this was addressed by applying thresholds to soil and land use classes that have a relatively small share within each subbasin (<10%). This limits the number of calculation units while at the same time preserving the heterogeneous nature of the basin and the different soi-landuse combinations that exist..

A summary of the SWAT main components relevant to this project is given in the following section.

2.2 Model specifications

2.2.1 Background

SWAT[1] was developed primarily by the United States Department of Agriculture (USDA) to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over long periods of time. The SWAT model has been extensively used, is in the public domain and can be considered as becoming the de-facto standard in spatial decision support systems.

SWAT represents all the components of the hydrological cycle including: rainfall, snow, snow-cover and snow-melt, interception storage, surface runoff, up to 10 soil storages, infiltration, evaporation, evapotranspiration, lateral flow, percolation, pond and reservoir water balances, shallow and deep aquifers, channel routing. It also includes irrigation from rivers, shallow and deep groundwater stores, ponds/reservoirs and rivers, transmission losses and irrigation onto the soil surface. It includes sediment production based on a modified version of the Universal Loss Equation and routing of sediments in river channels. SWAT has a modular set-up and it goes beyond the scope of this report to get into detail on each of these modules, but reference is made to the theoretical documentation (Neitsch et al, 2002).

For modelling purposes, a watershed may be partitioned into a number of sub-watersheds or sub-basins. Input information for each sub-basin is grouped or organized into the following categories: climate; hydrologic response units or HRUs; ponds/wetlands; groundwater; and the main channel, or reach, draining the sub-basin. Hydrologic response units are lumped land areas within the sub-basin that are comprised of unique land cover, soil, and management combinations. While individual fields with a specific land use, management and soil may be scattered throughout a sub basin, these areas are lumped together to form one HRU. HRUs are used in SWAT runs since they simplify a run by lumping all similar soil and land use areas within a sub-basin into a single response unit.

Simulation of the hydrology of a watershed can be separated into two major divisions. The first division is the land phase of the hydrologic cycle, depicted in Figure 1. The land phase of the hydrologic cycle controls the amount of water, sediment, nutrient and pesticide loadings to the main channel in each sub-basin.

[1] http://www.brc.tamus.edu/swat/index.html

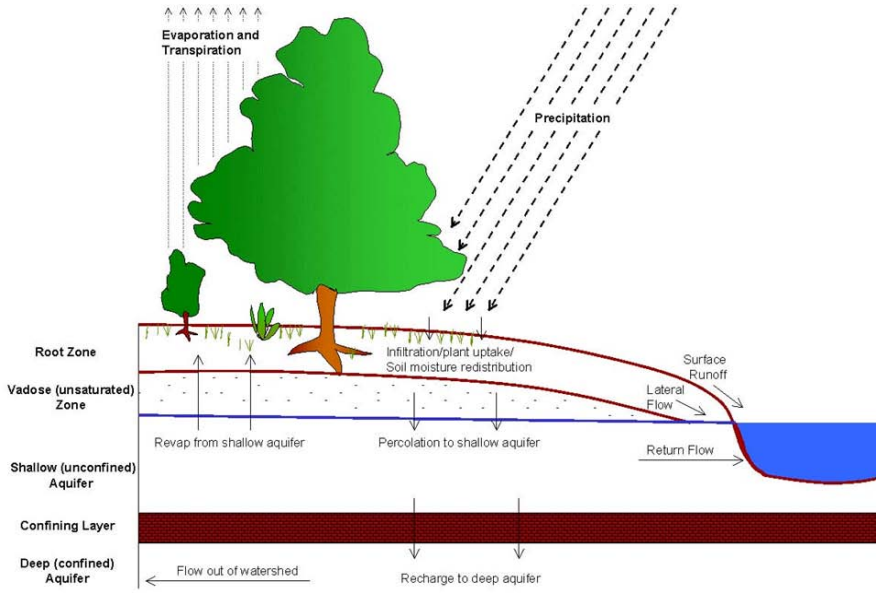

Figure 1: The water balance components as simulated by SWAT.

The second division is the water or routing phase of the hydrologic cycle which can be defined as the movement of water, sediments, etc. through the channel network of the watershed to the outlet. Once SWAT determines the loadings of water, sediment, nutrients and pesticides to the main channel, the loadings are routed through the stream network of the watershed using a command structure.

The most relevant components of the water balance and processes for this study are described in the following paragraphs.

2.2.2 Evapotranspiration

The model computes evaporation from soils and plants separately as described by Ritchie (1972). Potential soil water evaporation is estimated as a function of potential evapotranspiration and leaf area index (area of plant leaves relative to the area of the HRU). Actual soil water evaporation is estimated by using exponential functions of soil depth and water content. Plant transpiration is simulated as a linear function of potential evapotranspiration and leaf area index. Potential evapotranspiration is the rate at which evapotranspiration would occur from a large area completely and uniformly covered with growing vegetation which has access to an unlimited supply of soil water. This rate is assumed to be unaffected by micro-climatic processes such as advection or heat-storage effects. The model offers three options for estimating potential evapotranspiration. For this study the most advanced method will be used which is of Penman-Monteith (Penman, 1948).

2.2.3 Groundwater

Recharge to unconfined aquifers occurs via percolation of excessively wet root zones. Recharge to confined aquifers by percolation from the surface occurs only at the upstream end of the confined aquifer, where the geologic formation containing the aquifer is exposed at the earth's surface, flow is not confined, and a water table is present. River courses and irrigation canals are connected to the groundwater system, and surface water – groundwater interactions are taken care for.

After water is infiltrated into the soil, it can basically leave again the ground as lateral flow from the upper soil layer – which mimics a 2D flow domain in the unsaturated zone – or from return flow that leaves the shallow aquifer and drains into a nearby river. The remaining part of the soil moisture can feed into the deep aquifer, from where it can be pumped back by means of artificial extraction. The total return flow thus consists of surface runoff, lateral outflow from root zone and aquifer drainage to river.

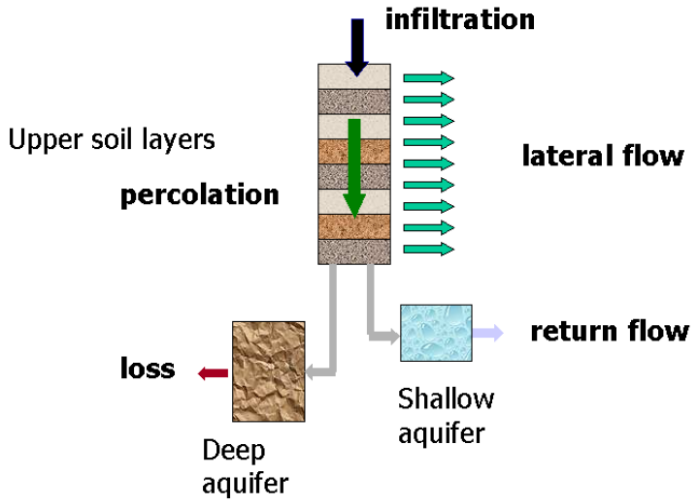

Figure 2: Schematic diagram of the partitioning of infiltration into sub-surface water fluxes after water uptake by roots have taken place

SWAT simulates two aquifers in each sub basin. The shallow aquifer is an unconfined aquifer that contributes to flow in the main channel or reach of the sub basin. The deep aquifer is a confined aquifer. Water that enters the deep aquifer is assumed to contribute to stream flow somewhere outside of the watershed (Arnold et al., 1993). The effects of groundwater extractions on base flow (Q_{gw}), defined as the contribution of the shallow aquifer to stream flow, is of specific relevance in this study.

2.2.4 Reservoirs

Reservoirs are located within a sub basin off the main channel. Water flowing into these water bodies must originate from the sub basin in which the water body is located. Reservoirs are located on the main channel network. They receive water from all sub basins upstream of the water body. A schematic representation of reservoirs in SWAT is shown in Figure 3.

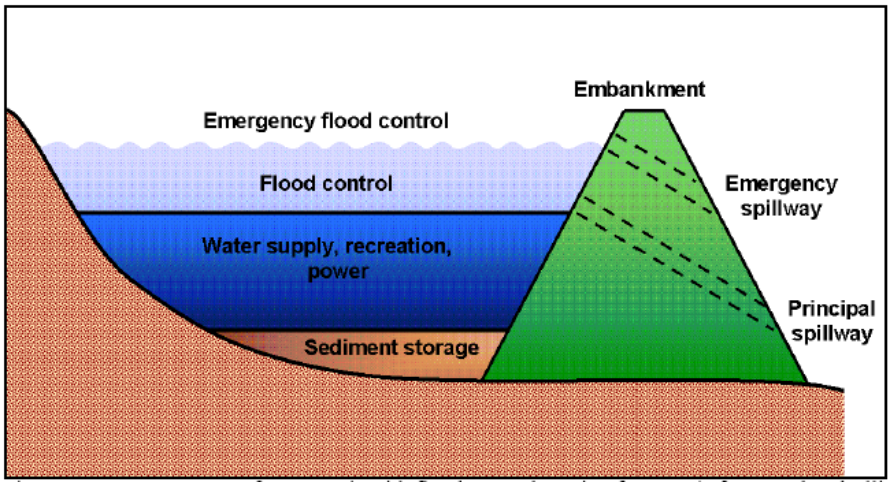

Figure 3: Schematic representation reservoirs in SWAT (Neitsch et al, 2001)

The water balance for reservoirs includes inflow, outflow, rainfall on the surface, evaporation, seepage from the reservoir bottom and diversions. Irrigation abstractions from the reservoir water storage can be simulated by SWAT, as is relevant for the Upper Tana model in this study.

2.2.5 Erosion

Erosion and sediment yield are estimated with the Modified Universal Soil Loss Equation (MUSLE) (Williams, 1975). While the USLE uses rainfall as an indicator of erosive energy, MUSLE uses the amount of runoff to simulate erosion and sediment yield. The substitution results in a number of benefits: the prediction accuracy of the model is increased, the need for a delivery ratio is eliminated, and single storm estimates of sediment yields can be calculated.

The rainfall-runoff scheme supplies estimates of runoff volume and peak runoff rate which, with the subbasin area, are used to calculate the runoff erosive energy variable. The crop management factor of the USLE equation is recalculated every day that runoff occurs. It is a function of above-ground biomass, residue on the soil surface, and the minimum C factor for the plant. Other factors of the erosion equation are evaluated as described by Wischmeier and Smith(1978).

2.2.6 Sediment routing

Sediment transport in the channel network is a function of two processes: degradation and deposition. These processes operate simultaneously in the reaches. The maximum amount of sediment transported through a reach segment is a function of the peak channel velocity. This maximum concentration is calculated as follows:

$$conc_{sed,ch,\max} = c_{sp} \cdot v_{ch,pk}^{sp\exp} \qquad \text{Eq. 2}$$

where $conc_{sed,ch,\max}$ is the maximum concentration transported that can be transported by the water (ton/m^3), c_{sp} is a coefficient defined by the user, $v_{ch,pk}$ is the channel peak velocity and *spexp* is an exponent defined by the user. The physically realistic ranges are known for both factors to be defined by the user in this equation. These ranges will be used to limit the calibration parameter space. This maximum concentration defines the amount of sediment deposited or reentrained in the reach segment and thus the amount of suspended sediment in the reach, which is a calibration variable for this baseline assessment model.

2.3 Model data requirements

An overview of the different steps to perform the hydrological and erosion modeling assessment using SWAT is provided in Figure 4. It also reveals the relation with the datasets that were used and generated using the tool.

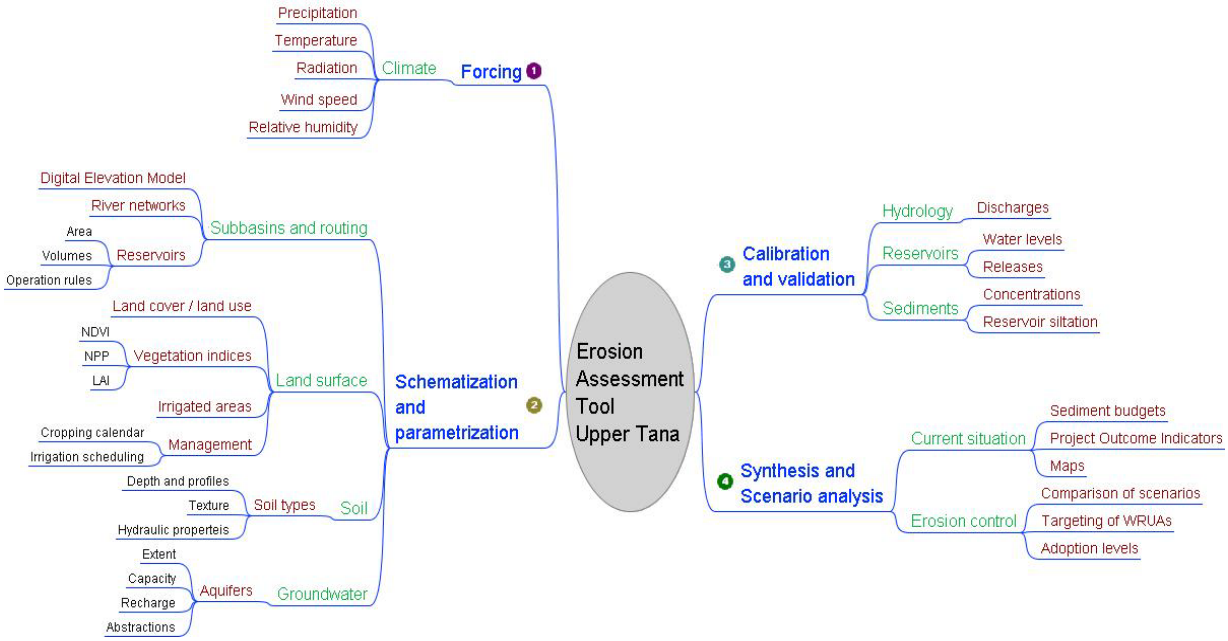

Figure 4. Diagram of used and produced datasets for the Upper Tana Erosion Assessment Tool.

Data was collected and analyzed from local databases provided by WRMA and previous studies and documents. In addition the remainder of the necessary data for the schematization of the model has been obtained from global public domain datasets.

The following sections describe the datasets gathered in the context of the soil erosion assessment model and procedures carried out to prepare them for the model input. The following datasets are described consecutively:

1. Climate
2. Digital elevation model
3. Soils
4. Land use and management
5. Streamflow datasets
6. Sediment and erosion datasets

2.4 Climate dataset

2.4.1 Required variables

The SWAT model requires meteorological data to be available at a daily time step. Besides precipitation, a crucial variable in the water balance is evapotranspiration. Several methods can be used to calculate the potential evapotranspiration in SWAT. The most comprehensive method available, which is the Penman-Monteith, requires data on temperature, solar radiation, wind and humidity for the calculation of the spatially distributed potential evapotranspiration rates.

In summary, SWAT requires the following variables for the Upper Tana model:
- daily rainfall
- minimum and maximum daily temperature

- solar radiation
- wind speed
- relative humidity

Daily data of these variables have to be available for the period 1980 – 2010, which is the period which will be used for calibration.

The Upper Tana watershed has a particularly accentuated orography, which causes a strong orographic precipitation regime. In general, rainfall amounts in the upper parts of Mount Kenya and the Aberdare Mountain Range are about 2 times the amounts in the lower regions. These strong meteorological gradients require an appropriate distributed approach for the rainfall input in the hydrological model.

A common problem in many basins is that meteorological stations are not available in the higher parts of the basin which are not well accessible. Only recently, satellite products have become available with sufficient resolution for hydrological studies to overcome this data gap. The reason why these freely available products are still not used very often, is that the processing of these datasets to make them useful for modelling studies requires advanced GIS and remote sensing techniques. At the same time, these products are the result of complex interpolation algorithms but normally have a certain amount of bias, different in each location. This makes it necessary to compare the estimates with gauged timeseries and correct them if necessary.

For the Upper Tana basin it was decided to use both gauged data as well as satellite rainfall estimates to obtain a high quality rainfall timeseries for each point in the basin. By using an advanced method to merge these two data sources, it was guaranteed that optimal use is made of the strengths of each dataset:

- The gauged weather station data because they are generally more precise in recording the absolute daily rainfall amounts
- Tthe satellite estimates which are more accurate in capturing the spatial rainfall patterns.

The following sections describe the datasets used and the procedures carried out to obtain daily gridded rainfall maps for the forcing of the simulation model.

2.4.2 Input variables from weather stations

The availability of weather stations in the Upper Tana basin is extensively described in the full project report (“Natural Resources Management Project: Consultancy Services For The Physiographical Survey In The Upper Tana Catchment, Loan No4772-KE”), Chapter 7. A major task of this project has been gathering the data itself of these stations, managed by the different institutes (WRMA, KMS, etc.). By the end of the contract, it has been possible to obtain a reasonable picture of the available data of all the stations. For the modeling activities to start in time and to be carried out successfully, it was decided to use standardized and validated data available from a global dataset with meteorological data. This way, the project outcomes of the modeling assessment did not depend on the parallel efforts that have been undertaken to request data from the local sources. Other advantages of using a global dataset are: (i) the dataset offers daily data, instead of monthly data of several stations that were available from local sources, (ii) the data has gone through a validation and standardization process, (iii) stations tend to be reliable and operational during long periods, assuring the coverage of at least a major part of the period required for model setup and calibration (in this case 30-year), and (iv).the datasets also include besides rainfall and temperature the other meteorological variables required (wind, dewpoint temperature).

The most useful global source for this study is the Global Summary Of the Day (GSOD) database archived by the National Climatic Data Center (NCDC). This database offers a substantial number of stations with long-term daily time series. The GSOD database submits all series (regardless of origin) to extensive automated quality control. Therefore, it can be considered a uniform and validated database where errors have been eliminated.

There are three relevant stations in this database that have been found useful for this study:(i) Embu, (ii) Nyeri and (iii) Nairobi Kenyatta Airport. Figure 5 shows the average monthly rainfall from these three stations, calculated from the daily time series from 1980 – 2010. The location of these stations is shown in Figure 6. The daily precipitation time series were used to generate daily gridded maps of the 30-year period using satellite rainfall distribution observations, as explained in the following sections. The temperature data was prepared as input for SWAT, in which a temperature lapse rate was assumed of 6 Celcius degrees per km. Relative humidity was derived from dewpoint temperature and solar radiation was calculated based on sunshine hours and geographic location for each of the stations.

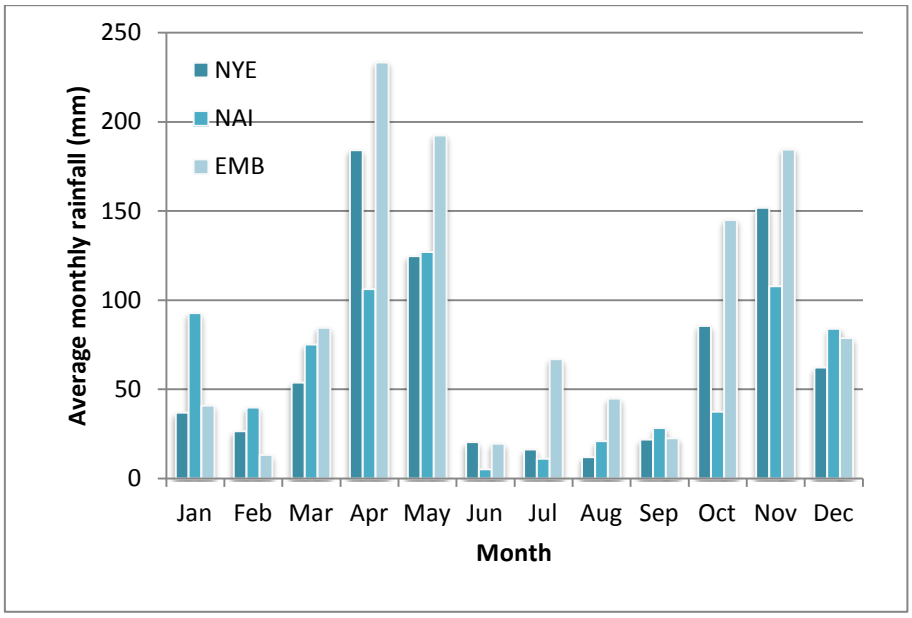

Figure 5. Average monthly rainfall from the Embu, Nyeri and Nairobi Kenyatta Airport weather stations

In the 30-years daily dataset (GSOD) some stations had a number of missing values. These missing values were filled by using the average value for precipitation on this particular day from the other stations. On average, the number of missing values per station was found to be about 15 per year in the simulated period 1980-2010.

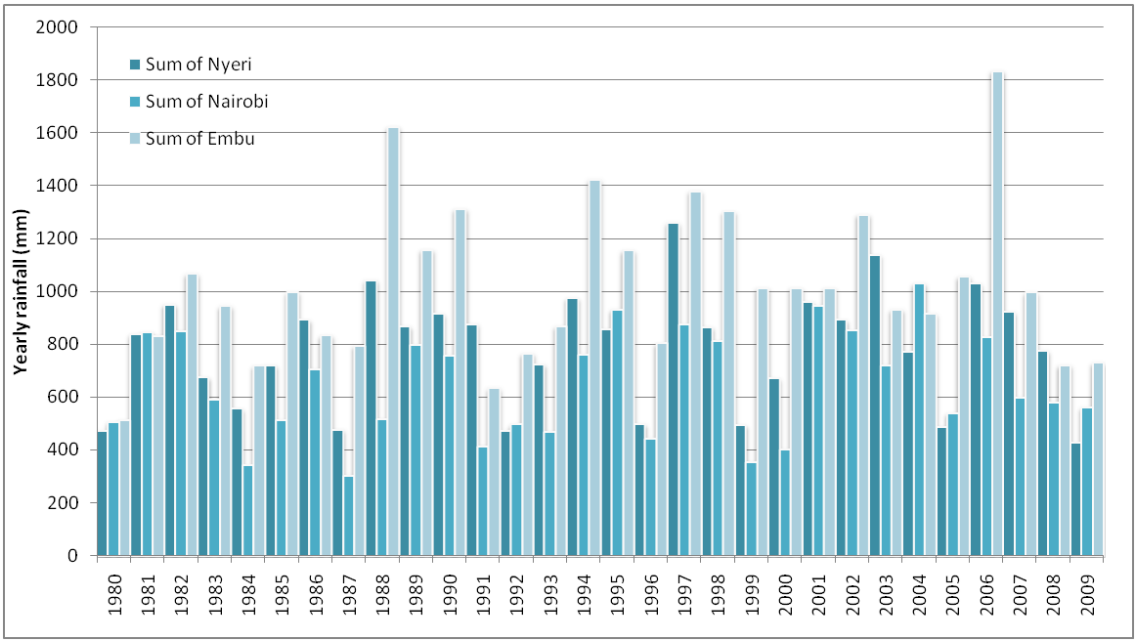

Figure 6. Total yearly accumulated rainfall for the 3 weather stations

For the years 1999 and 2000 the number of missing values was much higher for the Nyeri and Embu station: almost no data was recorded for these stations during these two years. Therefore, it was decided to fill these gaps by using the daily data obtained from WRMA recorded at the MIAD Climatic station near the Thiba river and very close to Embu. Outliers were detected by visual inspection and manually corrected.

2.4.3 Rainfall distribution from satellite

Many satellite rainfall products are available today, all of them with different characteristics, spatial and temporal resolutions, and coverages (f.e. PERSIANN, TRMM, CMORPH, TAMSAT). A good quality product focused on the African continent is prepared operationally at the Climate Prediction Center (CPC) for the United States Agency for International Development (USAID). They provide daily estimates of precipitation which are used for the Famine Early Warning System Network (FEWS NET). The algorithm for the rainfall estimates uses Meteosat 7 geostationary satellite infrared data that are acquired in 30-minute intervals, and areas depicting cloud top temperatures of less than 235K are used to estimate stratiform rainfall. This dataset has been found to outperform other satellite products in this part of Africa (Asadullah et al, 2008).

The latest product from FEWS NET is called RFE2 and is available from October 2000 until present with a spatial resolution of 0.1 degree (~10 km). This dataset has a certain site-dependent bias that has to be determined and corrected with ground-based observations (see also Hunink et al. 2009). Previous analysis for this part of Kenya found monthly mean absolute errors of 20 mm month^{-1}, and mean bias values of 15 mm month^{-1} (Funk and Verdin 2003).

For this particular study, this dataset does not provide a timeseries long enough to cover the entire simulation period (1980-2010). On the other hand, elevation, slope and orography in the area are of such importance that it is not recommendable to use only ground-based observations without taking into account the spatial rainfall patterns. For this reason, a

procedure was set up to enhance the daily ground-based observations with monthly information on rainfall distribution throughout the basin, based on FEWS RFE2.

2.4.4 Calculation of daily gridded rainfall input

To use the strengths of both data sources (gauged observations and satellite estimates), we have developed an analog to the satellite RFE2 estimates, enhancing the gauged time series with monthly averages based on the 10-year dataset of RFE2. Funk and Verdin (2010) use a similar method to derive improved gridded climatologies with satellite rainfall averages as a basis. This methodology benefits from the ability of satellite rainfall estimates to capture spatial gradients in rainfall and the relative accuracy of rainfall gauges.

Daily interpolated (inverse distance weighted) gauge-derived rainfall grids are enhanced by a correction grid that was derived for each month in the year, based on the daily 10 km resolution grids RFE2 of the years 2000-2010. These correction grids define the relative anomalies of each pixel compared to the closest gauge location. The first step to obtain these correction grids is to calculate for each month in the year the mean monthly-accumulated rainfall grids. The following step is to interpolate (inverse distance weighted) the values on the gauge locations of each of the twelve monthly mean grids. After, these interpolated grids are divided by the monthly mean grids themselves to obtain the relative anomalies of each pixel compared to the closest gauge location (being 0 at the gauge locations). In an equation, this can be stated as follows:

$$C = \frac{\langle P \rangle_{monthly}}{\langle P_{IDW} \rangle_{monthly}} \quad (1)$$

In which C is the correction grid, $\langle P \rangle_{monthly}$ = a grid for each month of the year representing the average rainfall of that month based on the daily RFE2 grids, and $\langle P_{IDW} \rangle_{monthly}$ = the interpolated (inverse distance weighted) grids of the RFE2 monthly averages at the weather station locations.

These monthly correction grids are used to obtain for each day a rainfall grid in which the rainfall amounts on the gauge locations are exactly the same as observed, while the interpolated values are based on the monthly mean grids. This procedure is summarized in the following equation:

$$P_{def,x} = C \cdot P_{IDW,x} \quad (1)$$

In which $P_{def,x}$ = the final used precipitation grid for the model input for day = x, $P_{IDW,x}$ is the interpolated (inverse distance weighted) precipitation grid based on the station timeseries and C is the correction grid for the month of this particular day x.

The methodology assures that the daily rainfall amounts on the station locations is exactly the same as those in the measured station timeseries. Monthly and yearly sums of the daily grids were evaluated and watershed averages were extracted. The following figure shows distribution of the total rainfall in 1984 within the watershed, calculated by summing all the daily grids. As expected, in the downstream part of the watershed falls less precipitation, and the highest amounts are found on the east-facing slopes of Mount Kenya and the Aberdares range.

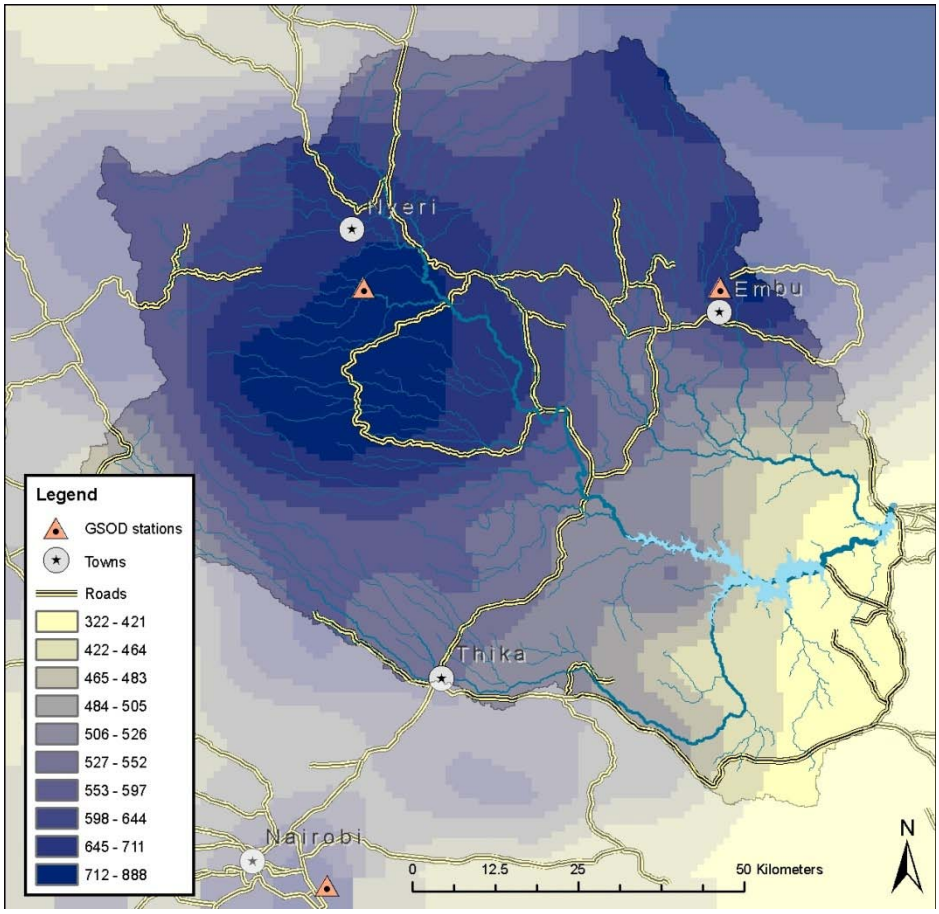

Figure 7. Example of the resulting spatial rainfall pattern for 2008, calculated as the sum of 365 daily grids, based on GSOD and FEWS data

2.5 Digital elevation model

Digital Elevation data are obtained from the Shuttle Radar Data Topography Mission (SRTM) of the NASA's Space Shuttle Endeavour flight on 11-22 February 2000. SRTM data were processed from raw radar echoes into digital elevation models at the Jet Propulsion Laboratory (JPL) in California.

SRTM data at 3 arc-second (90 meters) is available for global coverage between 60 degrees North and 56 degrees South latitude. The product consists of seamless raster data available in geographic coordinates (latitude/longitude) and is horizontally and vertically referenced to the EGM96 Geoid (NASA 1998). The SRTM-DEM data have been obtained using the USGS Seamless Data Distribution System (USGS 2004). Small voids present in the dataset within the area were filled by spatial interpolation.

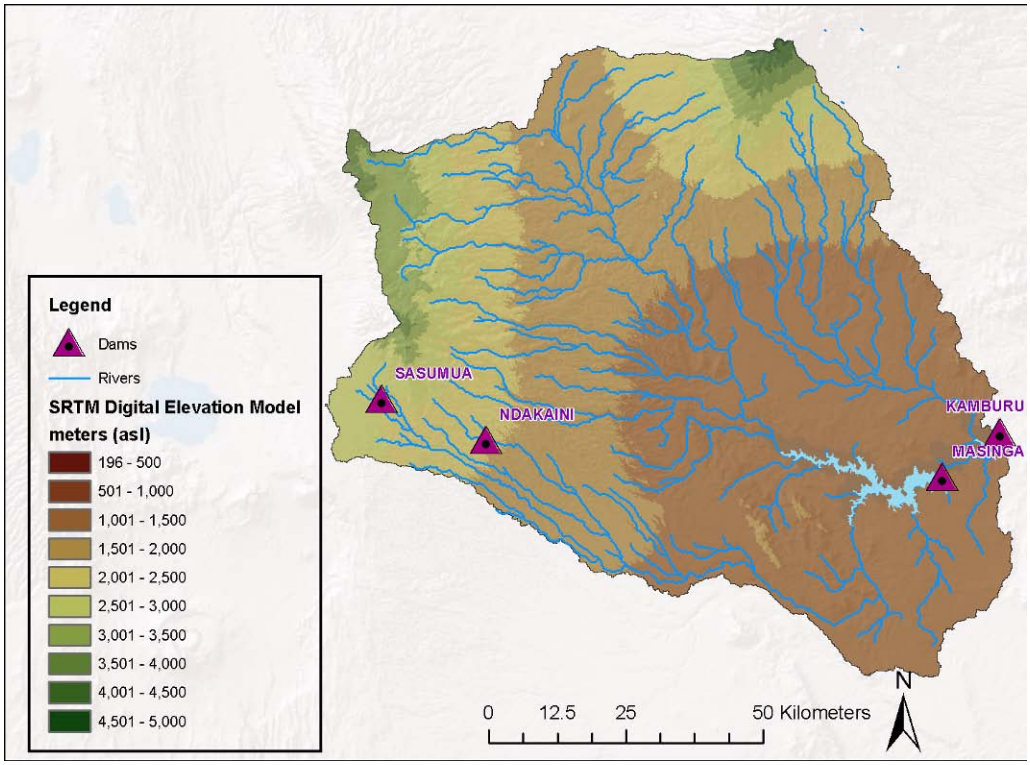

Figure 8. Digital elevation model used for network delineation

2.6 Soil dataset

2.6.1 Distribution of soils

The KenSOTER database at scale of 1:1 million (KSS 1996), holds data on landform, parent material and soils in a standardized digital format (van Engelen and Wen 1995). This database was updated by Kenya Soil Survey and ISRIC-World Soil Information (Batjes and Gicheru 2004). This 2004 version was expanded with additional profile data with measured water retention values of the Upper Tana catchment. The current KENSOTER database contains now data of 340 soil profiles, of which 68 of the Upper Tana, we will refer to it as the KenSOTER-version 2 database (KSS and ISRIC 2007).

The dominant soil types of the Upper Tana catchment are presented in Figure 9 and show a clear relationship with elevation. The higher slopes of Mt Kenya and the Aberdares are dominated by volcanic ash soils (Andosols). The middle foot slopes have mainly deep well structured nutrient rich clay soils (Nitisols). The lower foot slopes are dominated by very deep strongly leached poor clay soils (Ferralsols) and by less leached soils (Cambisols and Luvisols). At lower elevations, roughly below 1000m, Cambisols and sodic-alkaline soils (Solonetz) are the dominant soils (KSS 1996; Sombroek, Braun and van der Pouw 1982).

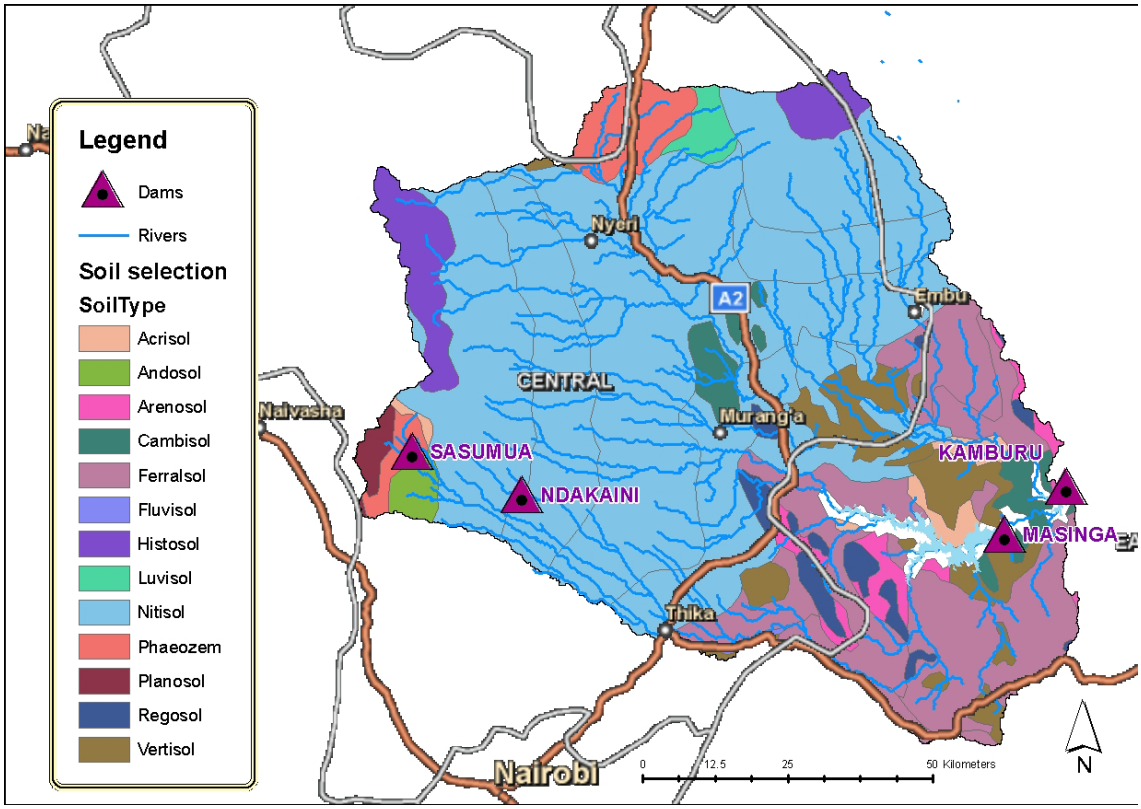

Figure 9: Dominant soil types of the Upper Tana catchment (KenSOTER-version 2)

2.6.2 Soil hydraulic parameters

The harmonized KENSOTER database is a secondary dataset with median attribute values. Missing entries are based on pedotransfer rules (van Engelen *et al.* 2005). Following these taxotransfer rules (Batjes 2003), the median attribute values have been estimated using attribute data and aggregate these over five fixed depth intervals, all on basis of texture group and soil unit classification. Soil classification follows the Revised Legend of the Soil Map of the World (FAO 1988).

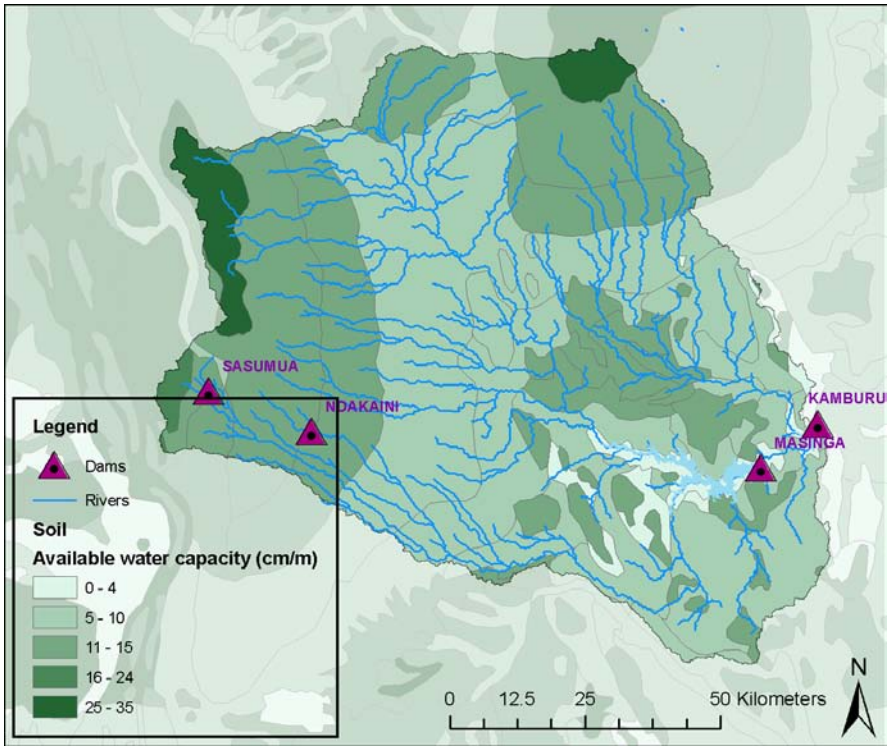

Figure 10: Available Water Capacity of dominant soils of the watershed (KenSOTER-version 2)

The harmonized KENSOTER database includes the total available water capacity of the soil, which data can be directly used in SWAT. The rootable soil depth is directly extracted from the harmonized KenSOTER database. In a few cases the rootable depths of the harmonized KENSOTER is somewhat different of KenSOTER-version 2, because of the use of different criteria.

Bulk density and Available Water Content[1] are key soil hydrological properties determining the water balance, which are used in SWAT. The geographic distribution and the differences are shown in Figure 10 and Figure 11. Comparing soil types in the Upper Tana it appears there is a factor 5 to 10 difference between lowest and highest values of Total Available Water Content.

[1] Available Water Content is the amount of moisture held between pF2.3 and pF4.2

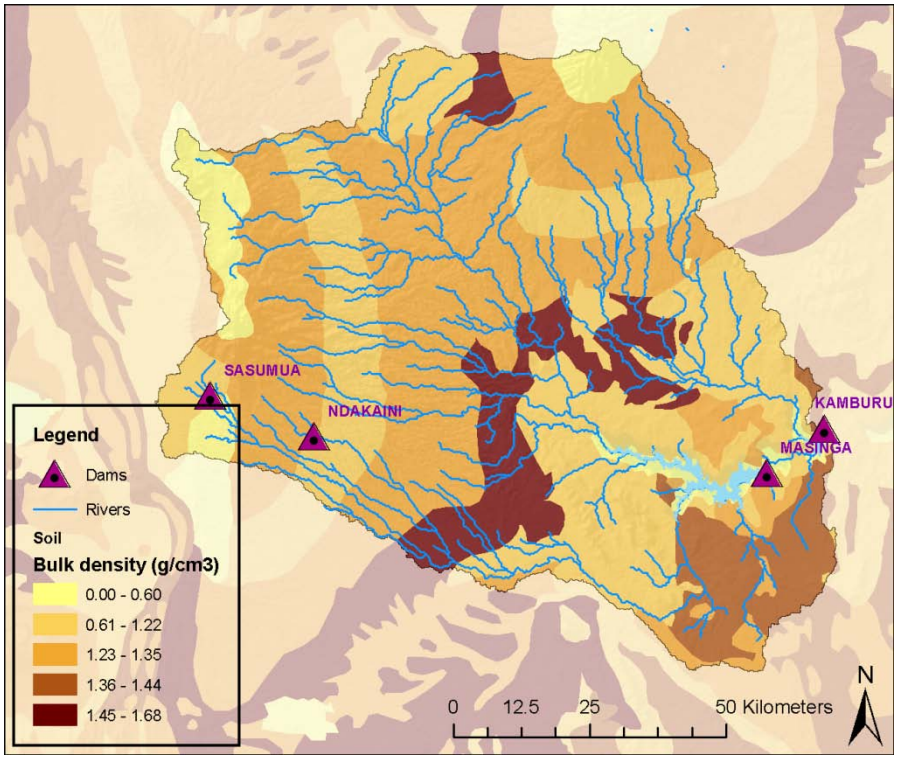

Figure 11: Bulk density of dominant soils of the watershed (KenSOTER-version 2)

The harmonized KENSOTER database contains most of the information necessary for the SWAT model.
An important characteristics not provided by KenSOTER database is the saturated hydraulic conductivity. A well-developed technique to overcome this problem is to use so-called pedo-transfer functions (PTF). A wide range of pedo-transfer functions have been developed and applied successfully over the last decades over various scales (e.g. field scale in (Droogers *et al.* 2001); basin scale at (Droogers and Kite 2001).

Sobierja et al.(2001) concluded from a detailed analysis that most PTFs were not very reliable and that the impact on runoff estimates could be considerable. The PTF that generated conductivity values close to measured ones was the Jabro equation (Jabro 1992):

$K_{sat} = \exp(11.86 - 0.81 \log(st) - 1.09 \log(cl) - 4.64\ BD)$

K_{sat} is saturated hydraulic conductivity (cm h^{-1})
st is silt content (%)
cl is clay content (%)

This equation was used to derive K_{sat} values for the SWAT model, based on the KENSOTER database.

2.7 Land use dataset

The best available maps for the Upper Tana of land cover and use are those from the FAO Africover project (FAO 2000) which designates land use/land cover for points on an approximately 2400 x 4800 m irregular grid. The effective scale is about 1: 250 000. The land

cover has been interpreted from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in the year 1999 and local information. The land cover classes have been developed using the FAO/UNEP international standard LCCS classification system.

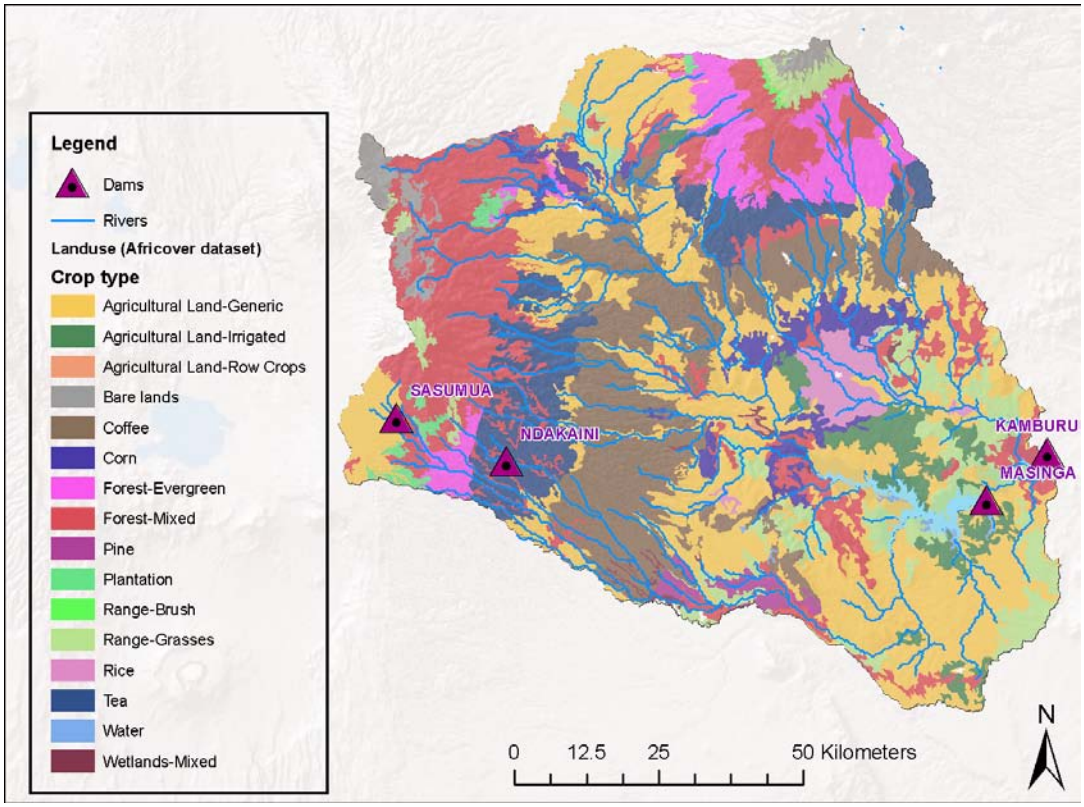

Figure 12: Landcover map as used in the SWAT model, main source: Africover dataset

The comparison of the Africover classified features with recent satellite imagery showed that the delimited features have not been altered significantly since the production of the dataset, taking into account the working scale of the study. However, a few polygons were misclassified and it was necessary to correct them, based on the visual comparison with the satellite imagery (publicly available through the NASA website). According to the original dataset these polygons had a dominating natural land cover but the imagery showed that the agricultural activities in those areas were more significant, especially for their hydrological relevance. Figure 12 shows the spatial distribution of the land covers as will be used in the SWAT model.

Table 1. Crop types and their relative importance in the watershed

Crop type	Area (km2)	% total
Agricultural Land-Irrigated	329	3.5%
Agricultural Land-Generic	2886	30.7%
Agricultural Land-Row Crops	5	0.1%
Bare lands	163	1.7%
Coffee	1491	15.8%
Corn	342	3.6%
Forest-Evergreen	544	5.8%
Forest-Mixed	1799	19.1%
Pineapple	78	0.8%
Plantation	72	0.8%
Rice	187	2.0%
Range-Brush	10	0.1%
Range-Grasses	700	7.4%
Tea	638	6.8%
Water	93	1.0%
Wetlands-Mixed	60	0.6%
Urban	14	0.2%
Total	**9,411**	**100.0%**

Table 1 shows how the land uses relate to the total basin area. As can be seen, a major part of the catchment is used for rainfed agriculture (principally maize, coffee and tea). Around 4% of the basin is used for irrigated agriculture (rice). Forest areas cover almost a third of the watershed area. Coffee is a relevant crop in terms of soil erosion as many fields often on steep slopes have been abandoned. Also maize cultivated on the steeper slopes of the Aberdares and Mount Kenya is thought to be a significant contributor to the basin soil erosion.

3 Model performance

3.1 Calibration procedure

SWAT is a complex hydrological model that requires knowledge and data on many processes and related parameters. Many of these parameters can be estimated or derived from catchment characteristics. Others have to calibrated in order to obtain a model that is able to simulate past conditions in an adequate way.

It is common practice to use observed flow discharge data to calibrate hydrological models. The complexity of the processes studied in this project requires a more elaborated calibration strategy as usual. To calibrate the model for this study, the following datasets are available: (i) historic flow discharge data from different measuring points in the basin, (ii) data on sediment loads from different locations that were monitored during this project, and (iii) new data on the long-term reservoir siltation rates. Using these datasets, the model calibration is carried out by going through the following consecutive steps:

1. Erosion rates and sediment yields depend mainly on rainfall intensity and runoff rates. The first step is therefore to use observed flow discharge data to calibrate the hydrological model. This will be based on monthly 30-year timeseries (1980-2009) calculated from daily measurements at various streamflow gauges throughout the basin.
2. The long-term calibration is validated with data inflow data from the Masinga reservoir that was calculated based on the water stages and the measured water outflow. A monthly dataset is available from 1982 – 2006.
3. The erosion rates and sediment yields are calibrated using the average siltation rate for each reservoir calculated from the bathymetric surveys. These rates will be compared with the model predictions to calibrate the soil erosion and sediment routing parameters of the model
4. The last step is to fine-tune the model with the data available from the current measurement campaign on streamflow and sediment loads in order to improve the spatial accuracy of the model parameters and the spatial heterogeneous output.

One important note that has to be taken into account is the following. The long-term calibration spans a 30-year time period in which the basin has changed considerably in terms of land use (for example increased extent of cultivated areas) and in terms of infrastructure (roads and small-scale hydraulic works, diversions, etc). For the calibration procedure it is assumed that all these factors are stationary and do not change over time. This is common practice in hydrological modelling but it is something to be aware of when interpreting the model response and the results.

3.2 Evaluation of calibration data

3.2.1 Streamflow data availability

A considerable amount of data on discharge measured at various stations throughout the basin has become available for this study. WRMA provided of different relevant stations the streamflow data that was obtained from 1980 onwards. **Table 2** provides an overview of the

gauging stations of which data is available for the long-term calibration of the erosion model. A few of these stations coincide with the locations monitored under this project, as indicated in the table. The locations of the stations are shown in **Figure 13**.

Table 2. List of available data points for calibration

Code	Watershed	Monitored under project	First date available	Last date available
4DA10	Thiba	No	01/01/1982	28/04/2005
4DC03	Thiba	No	01/01/1980	30/01/2007
4DD02	Thiba	Yes	01/01/1980	29/10/1993
4AC03	Tana (main)	No	01/01/1980	21/02/2010
4AD01	Tana (main)	No	01/01/1980	30/11/2008
4BC02	Tana (main)	No	01/01/1982	30/11/2008
4BE02	Tana (main)	no	01/01/1980	31/07/1981
4BE10	Tana (main)	yes	01/01/1980	21/02/2010
4AA05	Tana (tributaries)	no	01/01/1982	21/02/2010
4AB01	Tana (tributaries)	no	01/01/1980	27/02/2010
4BC05	Tana (tributaries)	no	01/01/1980	28/08/2003
4BD07	Tana (tributaries)	yes	01/01/1980	31/07/2003
4BE01	Tana (tributaries)	yes	01/01/1980	30/11/2008
4BE03	Tana (tributaries)	yes	01/01/1980	27/02/1997
4BE06	Tana (tributaries)	no	01/01/1980	28/05/1997
4BE08	Tana (tributaries)	yes	01/01/1980	30/08/1997
4BF01	Tana (tributaries)	yes	01/01/1980	21/02/2010
4CA02	Thika	no	01/01/1982	30/11/2008
4CB04	Thika	no	01/01/1982	21/02/2010
4CB05	Thika	yes	01/01/1980	29/03/2003
4CB07	Thika	no	01/01/1980	17/08/2003
4CC03	Thika	no	01/01/1980	28/12/1995
4CC07	Thika	yes	01/01/1982	21/02/2010

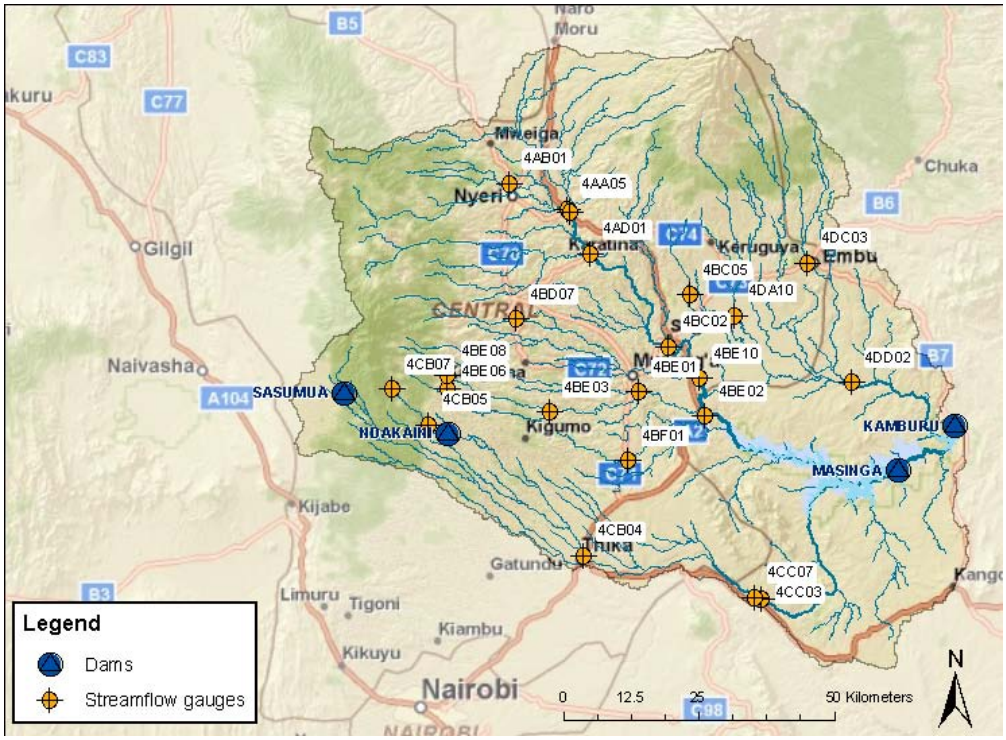

Figure 13. Map of dams and streamflow gauges of which data is available from 1980 onwards.

No quality flags were available for the measurements of the historic dataset nor other type of information on data quality, as for example past problems with the measurement instrumentation or interventions that might have changed the homogeneity of the timeseries. For this reason, the data quality assessment had to be fully done based on observations on the data itself as detailed in the following sections, while no past field observations or judgments from local staff could be included.

Consequently, the timeseries at the different locations were subjected to visual inspection and expert judgment. The final set of calibration points was based on these judgments. Within this context it has to be stressed that calibrating a model with erroneous data leads obviously to an erroneous model. During calibration, also in the final set of points a few issues were observed that raised doubts about the homogeneity and reliability of the timeseries. However, these issues were only in one case important enough to skip another point out of the long-term calibration procedure (4BD07).

In the following paragraphs, the suitability of each of the timeseries for model calibration will be evaluated, in order to obtain a final set of calibration points. For optimal output, it will be assured that the calibration points selected will cover each part of the basin and that each of the main catchments of the study area (Thiba, Upper Tana and Thika) includes several points.

3.2.2 *Streamflow data in Thiba catchment*

The Thiba catchment is draining into the Kamburu reservoir. Figure 14 shows the streamflow measured at the different monitoring locations during the period that will be used to calibrate streamflow (1980-2009).

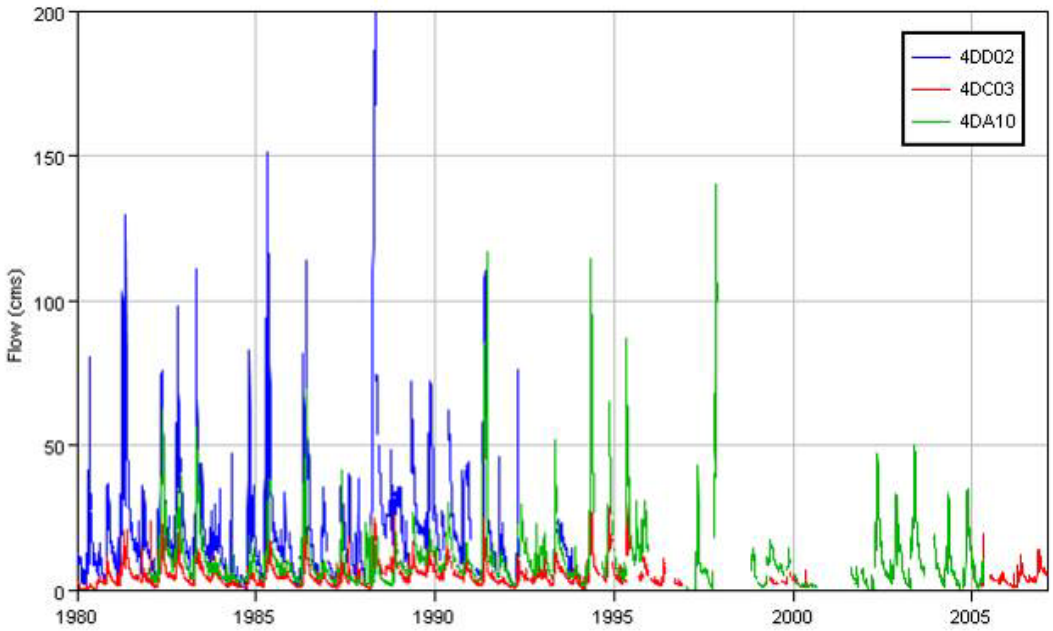

Figure 14. Streamflow plots of the gauges in the Thiba watershed

The following observations on the data of the Thiba catchment were done:

- The most downstream gauge 4DD02 (see **Figure 13**) has only data until 1994. This location is also currently being monitored within this project. However there is a data gap from 1995 to 2010. In spite of the relatively poor record, this station will be included in the long-term calibration procedure due to its importance and location.
- The two upstream gauges available (4DA10 and 4DC03) can partly fill the data gap, as becomes clear from Figure 14. Although they are located more upstream, these datasets contain valuable information on the years which are not covered by 4DD02. and will therefore be included in the long-term calibration procedure to ensure a reliable calibration for the Thiba catchment.

In summary, a total number of 3 stations will be used for the long-term calibration of the Thika catchment, being: 4DD02, 4DA10 and 4DC03.

3.2.3 *Streamflow data in Tana catchment*

The Tana catchment upstream Masinga reservoir has many gauges on the main stem, which are plotted in Figure 15. Also data of several gauges on the tributaries have become available to complement data deficiencies if necessary.

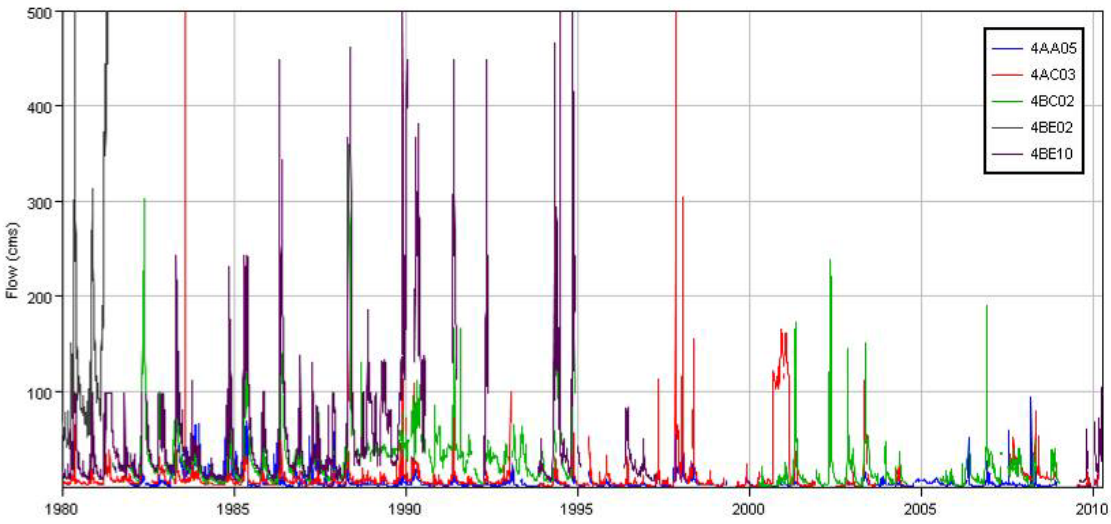

Figure 15. Streamflow of the gauges on the main stem of Tana river

The following observations can be done on the data of these gauges:

- The most downstream gauge 4BE02 has too few data available that fall within the calibration period (1980 onwards). There are only daily measurements for less than two years (1980 and 1981). For this reason, this station is omitted from the calibration procedure.
- Another downstream gauge of this catchment is 4BE10 and is currently monitored under this project. The historical record of this station shows a few problems. First, from 1980 until 1983 the data seems to be cut off at 100 m3/s. From 1983 onwards, there are gaps and there seem to be larger floods that are not registered on the upstream stations on the Sagana. Due to these data quality problems, this station is omitted from the calibration procedure.
- The three Sagana main stem stations with sufficient records, 4AC03, 4BC02 and 4AA05, seem to correlate with each other although they are not always proportional in flood magnitude, which can be caused by variable rainfall patterns in the area. The fact that they show a similar behavior makes them candidates for calibration of the main Tana stem to compensate the deficiencies of the downstream gauge 4BE10. The two stations with the largest upstream area are selected for the calibration procedure (4AC03 and 4BC02)
- The station 4BD07 is currently measured under this project. From the 30-year record, a mean flow of about 10m3/s can be derived. This is a relatively very high flow compared to the small draining area (about 125 km2). Recalculating, this would imply an average water yield of about 2000 mm. This is unrealistically high and about 3 or 4 times higher then the water yields of nearby stations. Also the measurements done during this campaign show too high values compared to the drained area. This station is omitted from the long-term calibration procedure. The issue will be further inspected during the last phase of the project.
- Three other stations with sufficient data of the tributaries of this catchment are 4BD07, 4BE03 and 4BE01. No major problems have been observed with these data so they were included in the calibration procedure.

3.2.4 *Streamflow data in Thika catchment*

The Thika catchment drains into the Masinga reservoir. It contains two smaller reservoirs managed by Nairobi Water Supply which are included in this study: Sasumua and Ndakaini. The following figure shows the streamflow plotted from 1980 onwards of the main stations of this catchment.

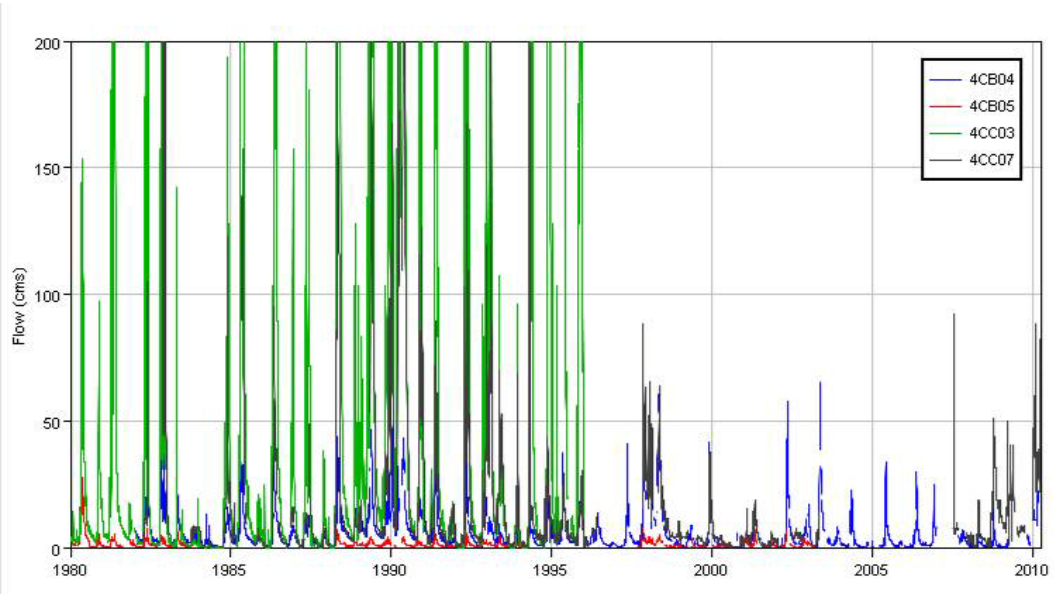

Figure 16. Streamflow of the gauges on the main stem of Thika river

The following observations have been done on the historic streamflow data of the Thika catchment:

- The most downstream gauge of the catchment (4CC07) shows erroneous peak flow measurements before 1996, which becomes obvious when looking at **Figure 16** (note that the most of the peak flows are around 700 m3/s) These extreme flows, also recorded for the close station 4CC03 (Yatta Furrow) do not relate with the measurements after 1996 and neither with the flow rates currently measured (being a wet year). It is likely that a rating curve error is the cause of this invalid data points. This station is therefore not suitable for calibration.
- A good candidate to compensate the data deficiencies for this catchment is 4CB04 (above the confluence with Chania river) that has a reliable continuous timeseries that can be useful for calibration of the Thika watershed for the period before 1996.
- Another station with a sufficient historical record is 4CB05. It drains a relatively small area which in principle makes it a less valuable station for calibration. Nevertheless, as there are no other suitable stations in the Thika catchment, this station is included in the long-term calibration procedure.

Part of the water retained in the NWS reservoirs is diverted outside of the catchment. In the model it was assumed that 360,000 m^3/day is diverted from the Ndakaini Dam and 60,000 m^3/day from Sasumua Dam (extracted from UN Kenya National Water Development Report, 2006).

3.2.5 *Streamflow data for long-term calibration*

Based on the evaluation of the timeseries from 1980 onwards from all the locations available, a selection was made for the long-term calibration. As discussed before, it was decided to use a

relatively high number of complementary stations to fill the data gaps, due to data deficiencies of the downstream stations. These additional stations are located more upstream but all together they cover a major part of the basin (see **Figure 17**).

The key objective is to have reliable estimates of stream and sediment inflow into the 4 reservoirs and to have an overall good performance downstream. To check whether this is the case, it was decided to validate the model with one additional dataset derived from the water stages in the Masinga reservoir and the measured outflow. Based on the water balance, the inflow can be derived with relative high accuracy on a monthly time basis. This timeseries was used to validate the overall performance of the model, using different performance criterions.

Table 3. List of locations that will be used for calibration

Code	Catchment	Calibrate	Reason
4AC03	Tana (main)	✓	OK
4AD01	Tana (main)	✓	OK
4BC02	Tana (main)	✓	OK
4BE02	Tana (main)	✗	Too many data gaps
4BE10	Tana (main)	✗	Unconsistent dataset, as explained in text
4AA05	Tana (tributaries)	✗	Covered by downstream station 4AC03
4AB01	Tana (tributaries)	✗	Covered by downstream station 4AC03
4BC05	Tana (tributaries)	✗	Covered by downstream station 4DD02
4BD07	Tana (tributaries)	✗	Unconsistent dataset, as explained in text
4BDNew	Tana (tributaries)	✗	No historic data
4BE01	Tana (tributaries)	✓	OK
4BE03	Tana (tributaries)	✓	OK
4BE06	Tana (tributaries)	✗	Covered by downstream station 4BE03
4BE08	Tana (tributaries)	✗	Covered by downstream station 4BE03
4BF01	Tana (tributaries)	✗	Too many data gaps
GI1	Tana (tributaries)	✗	No historic data
4DA10	Thiba	✓	OK
4DC03	Thiba	✓	OK
4DD02	Thiba	✓	OK
4CB04	Thika	✓	OK
4CB05	Thika	✓	OK
4CB07	Thika	✗	Covered by downstream station 4CB05
4CC03	Thika	✗	Unconsistent dataset, as explained in text
4CC07	Thika	✗	Unconsistent dataset, as explained in text
SA1	Thika	✗	No historic data
TH2	Thika	✗	No historic data

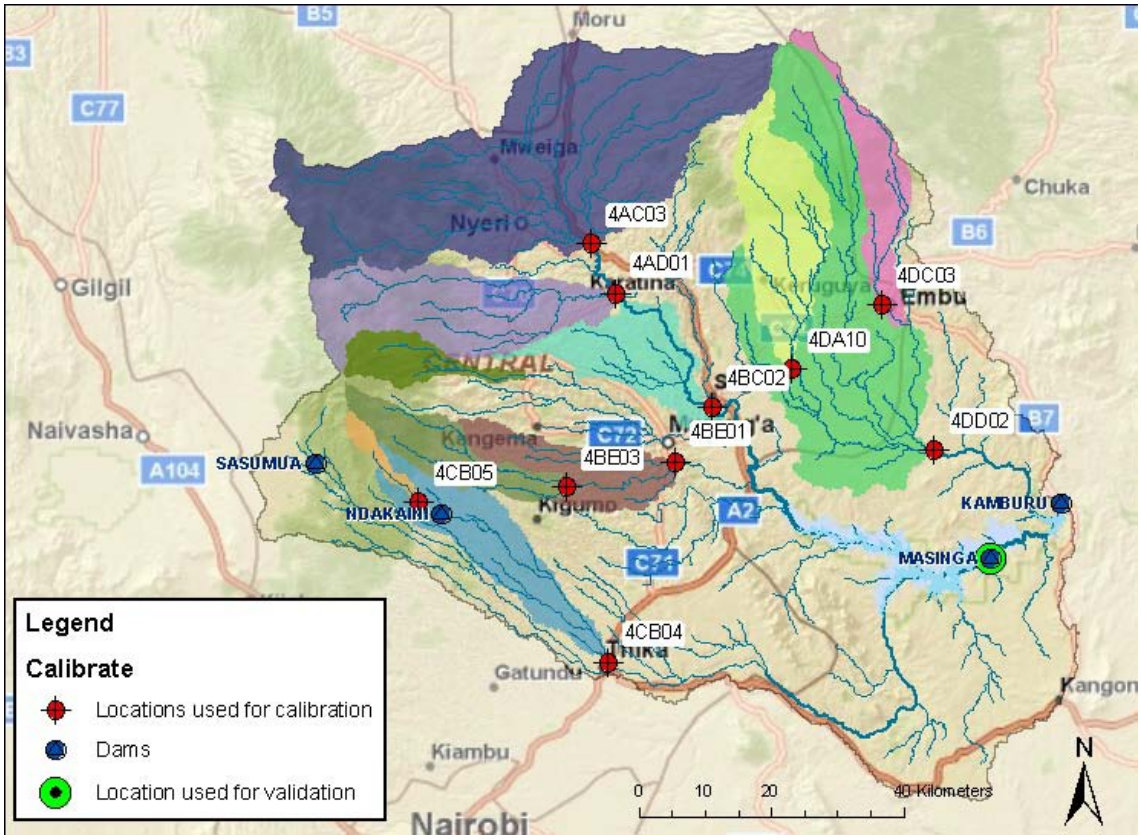

Figure 17. Locations used for calibration and validation with in different colors the corresponding upstream area with each calibration point

3.2.6 Reservoir sedimentation rate

The procedure used to derive the reservoir sedimentation rates, the gross sediment inflow into the reservoirs and the sediment budget for the reservoirs (Masinga and Kamburu) are presented in detail in Section 6.6 of Chapter 6 of the Full Report (see Preface). Beginning with the results of the bathymetric surveys which provided an accurate estimate of the reservoir full volumes, the sedimetation rates for the reservoirs were derived after comparing the present volume with a reference reservoir volume in the past.

For Masinga this resulted in a reservoir sedimentation rate of 5.45 Mcm/year, meaning the reservoir lost about 10.1% of its capacity over 29 years. For Kamburu, a sedimentation rate of 0.70 Mcm/year was derived (for the period after Masinga was constructed). The Kamburu reservoir is smaller and therefore has a larger problem of siltation: it lost about 14.7% of its capacity since 1983 (27 years). Kamburu reservoir had lost another 10% of its initial capacity between 1974-1983.

After accounting for both the effect of consolidation of sediments in the reservoirs with the passage of time, as well as for the trap efficiency of each reservoir (estimated analytically per sediment fraction of the mix), the sediment budget of the reservoirs and the gross sedment inflow rates from the upstream catchments were estimated. Masinga receives a yearly average of 8.03 Mton/yr of sediment inflow while Kamburu receives a further 1.1 Mton/yr influx (from the Thiba catchment) while Masinga outflows contribute another 1.28 Mton/yr. The sediment budgets of the reservoirs are described in detail in Section 6.6 of Chapter 6 of the Full Report.

The sediment input enters the reservoirs both as bed load as well as suspended load. By far most of the sediment enters the reservoir suspended in the water. The study finished in 2010 on Micro Projects Investments in the Upper Tana catchment (WRMA, 2010) summarizes previous studies (see Table 5.4) in the area in which estimations were done on suspended and bed load. The same report on Micro Projects Investments in the Upper Tana Catchment (WRMA, 2010) estimated also the contribution of roads. According to estimations done by the consultants that elaborated this study, about 25% of the sediment yield can be contributed to uncontrolled run-off along roads, loosened earthworks and culvert discharge into unprotected lands. Besides, they estimated that about 10% of total sediment yield is attributed to run-off generated from urban centres, institutions and homesteads without rainwater harvesting structures or soil and water conservation systems. The model does not incorporate these artificial works which tend to be very local in space and time, but the results will quantify their importance by comparing the simulated values of sediment inflow with the calculated values from the bathymetric surveys.

Table 4 Sediment load estimates from previous studies (based on WRMA, 2010).

Authors	Suspended sediment (tons/year)	Bed load (tons/year)	Total sediment load
1. Gibb (1959); *1947 – 1958 Q Records*		280,000	N/A
2. ILACO (1971); *1948 – 1965 Q Records*		250,000	N/A
3. Dunne (1975); *1956 – 1970 Q Records*	2,924,638	1,283,000	4,207,638
4. Dunne and Ongwenyi (1976); *1956 – 1970 Q Records*	5,275,638	568,547	5,844,185
5. Ongwenyi (1978); *1956 – 1970 Q Records*	4,580,000	519,000	5,099,000
6. Edwards (1979); *1956 – 1970 Q Records*	1,089,270	334,730	1,424,000
7. Otieno F.A.O. & Maingi S.M. (1984)	8,460,000	850,000	9,310,000
8. Maingi (1991); *1970 – 1984 Q Records*	7,470,000		N/A

3.3 Model calibration and validation

3.3.1 Long-term streamflow calibration

The first step in the calibration of the erosion model is to calibrate with the 30-year streamflow data scattered of the selected locations scattered throughout the basin (see also **Figure 17**). A sensitivity analysis was carried out to evaluate which of the possible calibration parameters are most responsive for the ratio fast runoff - baseflow.

Two parameters were adjusted on the entire basin scale and set to fixed values for each subbasin (Alpha_BF and Gw_Delay, see **Table 4**). Another parameter (SLSOIL) was adjusted using a basin-wide multiplier to improve the ratio lateral interflow and baseflow. By increasing this parameter, the travel time for lateral flow through the soil layers becomes higher and relatively more water is available for percolation to the aquifers. The final value is different for each subbasin (376 total).

In order to further improve the model parameterization and account for the large spatial heterogeneity in the basin, one additional calibration parameter was included in the procedure. This parameter was adjusted separately for each of the upstream areas of the calibration points (as shown in different colors in **Figure 17**). The parameter used for this distributed procedure is called RFINC which allows adjustments of daily rainfall (defined for each month of the year). Inaccuracies in the baseline data on land use (determining evapotranspiration), daily rainfall and

other meteorological input variables are accounted for by fine-tuning this parameter. This approach assures that the model response of each upstream part of the gauging locations is similar to the observed response.

The following table summarizes the adjustments made to the calibration parameters:

Table 5. Parameters used for streamflow calibration, initial and final ranges

SWAT Code	Unit	Description	Initial range	Final value
Alpha_BF	1/days	Baseflow recession constant	0.01 – 0.1	0.015
GW_DELAY	days	Groundwater delay time	15 – 60	50
SLSOIL	m	Slope length for lateral subsurface flow	10 – 2000	90 – 1200
RFINC(month)	%	Daily rainfall adjustment	0 – 300	0 – 190

Of each of the calibration points, the observed monthly response was compared with the simulated response and the correspondence quantified with two performance coefficients. The first is the the Pearson product-moment correlation coefficient and the second the Normalized Root Mean Square (RMS). The Normalized Root Mean Squared is the RMS divided by the maximum difference in the observed streamflow values, and is expressed by the following equation:

$$NormalizedRMS = \frac{RMS}{(X_{obs})_{\max} - (X_{obs})_{\min}}$$

The Normalized RMS is expressed as a percentage, and is a more representative measure of the fit than the standard RMS, as it accounts for the scale of the potential range of data values. For example, an RMS value of 1.5 will indicate a poor calibration for a model with a range of observed values between 10 and 20, but it will indicate an excellent calibration for a model with a range of observed values between 100 and 200. The Normalized RMS value for the first model would be 15%, while the Normalized RMS for the second model would be 1.5%.

The following table shows the two performance coefficients after calibration. The Pearson coefficient is around 0.7 and the normalized RMSE indicates a relative error between 8% and 21% depending on the location. Overall, the normalized RMSE, being a good indicator for the model uncertainty, is about 12%. The overall model performance is further assessed using the Nash-Sutcliffe criterion with the validation dataset.

	4AC03	4AD01	4DC03	4DA10	4BC02	4DD02	4BE01	4BE03	4CB05	4CB04	Total
Pearson coefficient	0.7	0.8	0.7	0.7	0.6	0.8	0.8	0.8	0.6	0.7	**0.7**
Normalized RMSE	8%	13%	14%	8%	13%	9%	11%	13%	21%	12%	**12%**

When plotting the observed monthly response against the simulated response, a few observations can be done and conclusions can be drawn. **Figure 18** shows the monthly streamflow measured and simulated at one of the tributaries of the Tana river. In general, both timeseries show a very good correspondence. Only a few peaks were not registered by the model, most probably due to local storms that were not recorded by the used rainfall stations.

Both low as well as high flows show good agreement (see for example 1997 – 1998). The normalized RMSE is 11%.

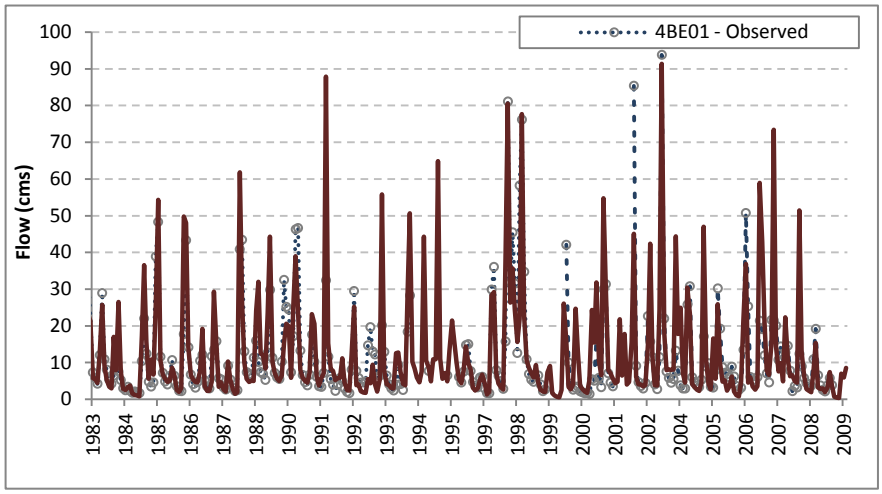

Figure 18. Observed versus simulated response for location 4BE01

Figure 19 shows both the observed and the simulated time series for a location on the main branch of the Upper Tana. The agreement between both time series is reasonably well (Normalized RMSE of 13%). However, it seems that there has been a significant change in hydrologic response of the upstream catchment during the entire period. Very few low flows are registered before 1994, with an average flow of about 35 m3/s. Between 1994 and 2000, no data was recorded at this location. After the year 2000, average flow is less than half of what it was before: 14 m3/s.

This inconsistency may be caused by different factors. A very likely important cause is gradual or abrupt upstream land use and management change: a process that affects either positively or negatively the upstream water balance (more or less loss through evapotranspiration). In the 30 years the model is representing, the basin has been going through many changes due to population growth, forestry management, farmers going to upstream areas, small hydraulic infrastructure, groundwater extractions, etc. These processes affect the hydrologic behavior of the catchment. This example illustrates that the model simulates well the average long-term catchment's conditions and response while at the same time it is inevitable to observe dissimilarities at certain locations due to gradual and abrupt upstream changes in the past.

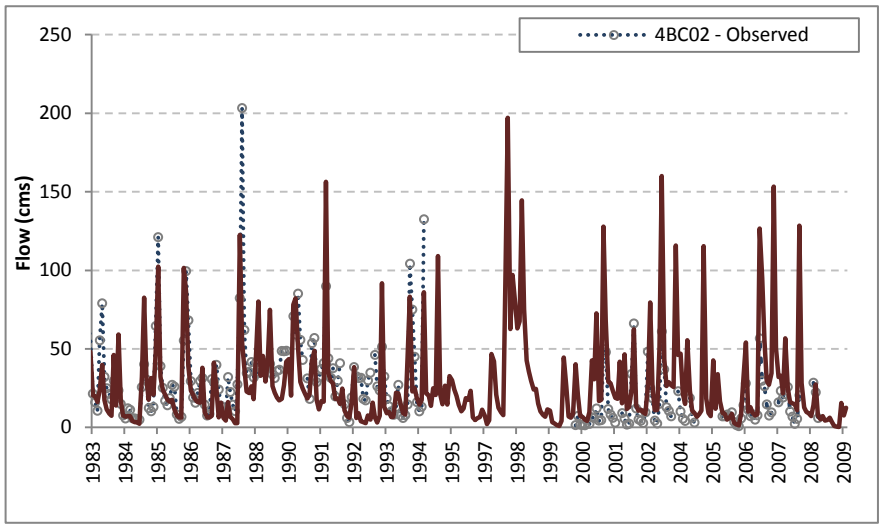

Figure 19. Observed versus simulated response for location 4BC02

The observed time series show large differences between high and low flows during the wet and dry seasons. Most of the sediment is transported during the wetter seasons. For this reason, it is important to assess whether the model correctly imitates the dynamic response during the year.

Figure 20 shows the average monthly response calculated from the observed and the simulated time series for two locations used for calibration. The same figures of the other locations are included in the Appendix of the report. As can be seen, the overall response of these two locations is fairly good. Also of the other used locations, a good agreement has been found between observed and simulated monthly response (R^2 = 0.9, see **Figure 21**). This demonstrates that the model performs well both under dry as under wet conditions giving confidence to the following calibration step using sediment data.

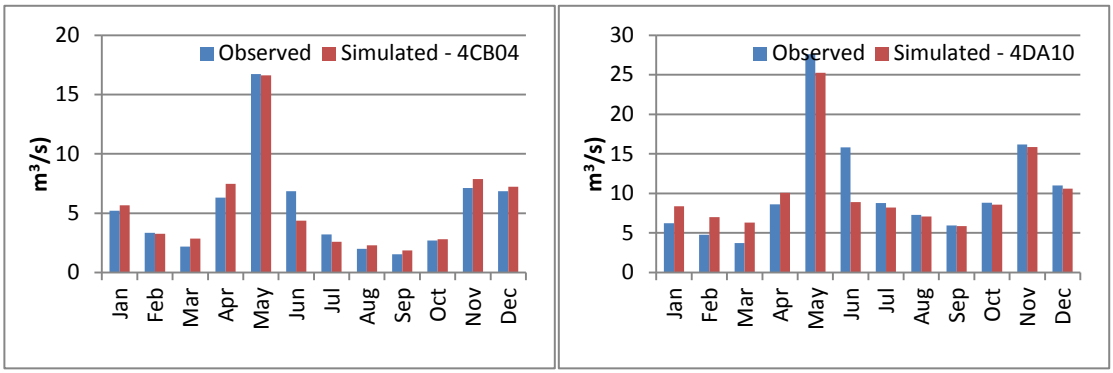

Figure 20 a and b. Monthly average observed and simulated streamflow for two locations

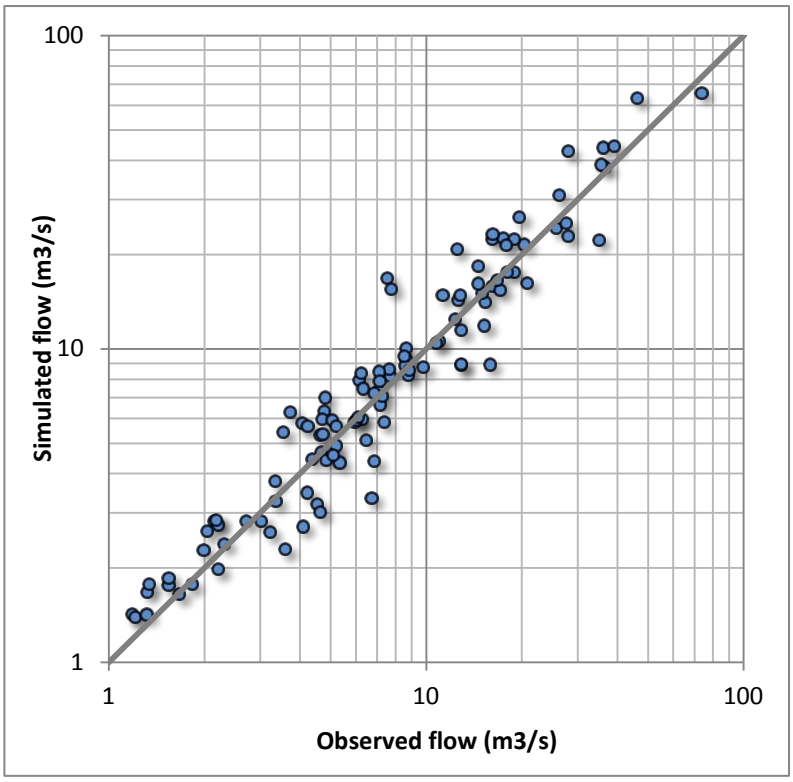

Figure 21. Observed monthly averages against simulated monthly averages of all calibration locations

3.3.2 Long-term streamflow validation

In order to further validate the model before calibrating the erosion and sediment yields, one additional dataset derived from water stages of the Masinga reservoir was included in the model evaluation. It has to be noted that this dataset was not used to calibrate the model, in order to make it a valid and fair validation dataset. This dataset is especially suitable for validation as it represents the major part of the total basin response (except of Thiba catchment).

The timeseries represents the total water inflow into the Masinga reservoir and was based on measured reservoir outflow and monthly average stages of the reservoir water levels. By making a mass balance of the reservoir, the net water inflow can be derived. The timeseries is available for 1983 until 2006, which means that it cover almost the entire calibration period.

The performance of the validation dataset was evaluated with the Pearson coefficient, the normalized RMSE (as described before) and the Nash–Sutcliffe model efficiency coefficient. This coefficient is a common indicator for overall model performance, and is defined as follows:

$$E = 1 - \frac{\sum_{t=1}^{T} (Q_o^t - Q_m^t)^2}{\sum_{t=1}^{T} (Q_o^t - \overline{Q_o})^2}$$

where Q_o is observed discharge, and Q_m is modeled discharge. Q_o^t is observed discharge at time t.

Nash–Sutcliffe efficiencies can range from −∞ to 1. An efficiency of 1 (E = 1) corresponds to a perfect match of modeled discharge to the observed data. An efficiency of 0 (E = 0) indicates that the model predictions are as accurate as the mean of the observed data, whereas an efficiency less than zero (E < 0) occurs when the observed mean is a better predictor than the model.

Table 6. Model performance indicators of validation dataset

	Pearson coefficient	Normalized RMSE	Nash Sutcliffe coefficient
Validation Masinga Inflow	0.87	9%	0.75

The overall good performance can be further observed from **Figure 22** and **Figure 23**. The first figure shows both timeseries over time in which can be seen that both high flows as well as low flows compare very well. Also the scatterplot shows that the points are all located around the x=y imaginary line.

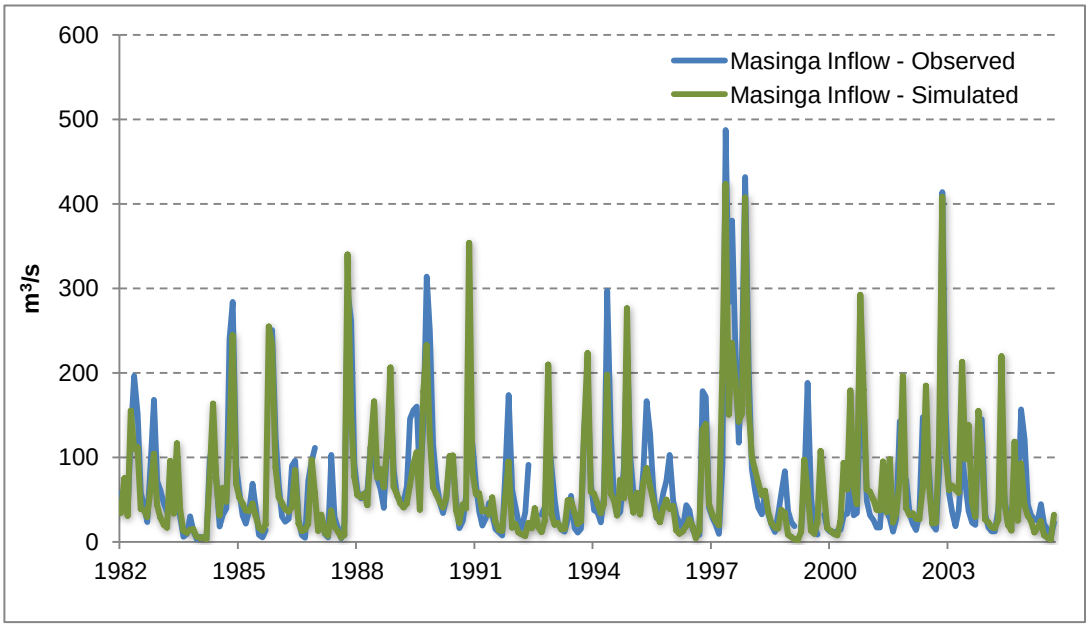

Figure 22. Timeseries of observed versus simulated water inflow of Masinga reservoir

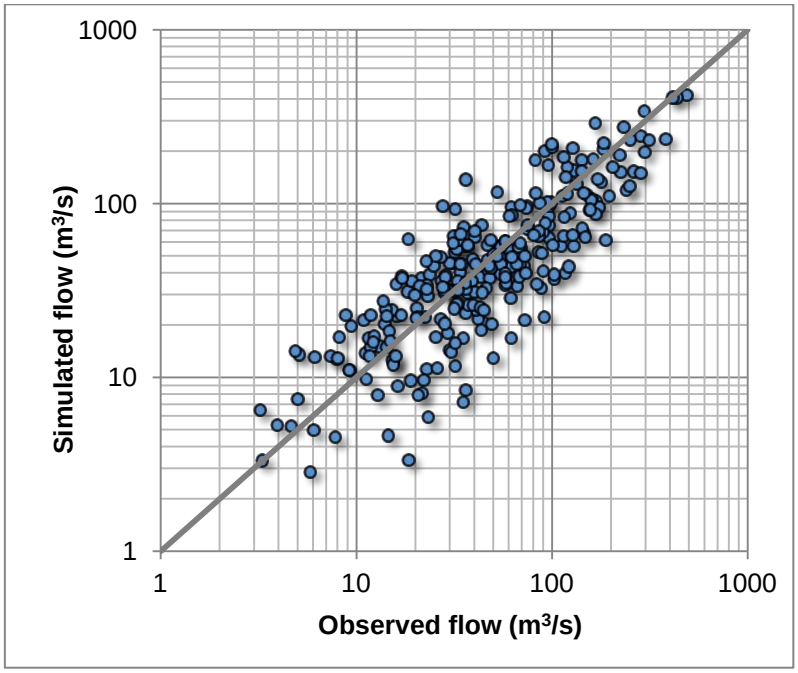

Figure 23. Scatterplot of observed versus simulated water inflow of Masinga reservoir

3.3.3 Erosion and sediment calibration

For the calibration of the erosion rates and sediment budgets, the following data was used:

- The derived long-term sedimentation rates from the bathymetric surveys
- The ratios between the average sediment loads at the different locations measured during the project
- The sediment discharges based on the discharge and sediment load measurements during project monitoring campaign.

The transport of sediment in the channel is controlled by the simultaneous operation of two processes, deposition and degradation. SWAT uses Williams (1980) simplified version of Bagnold's definition of stream power to develop a method for determining degradation as a function of channel slope and velocity. The maximum amount of sediment that can be transported from a reach segment is simulated as a function of the peak channel velocity. Available stream power is used to reentrain loose and deposited material until all of the material is removed. The capacity to reentrain the material depends on two parameters, which were calibrated for the basin as a whole (see **Table 6**).

To further improve the relative contribution of each of the sub-catchments of the basin to the total sediment production, it is necessary to use a distributed calibration parameter. The soil erodibility K factor was fine-tuned (**Table 6**) based on the sediment budgets calculated from the bathymetric surveys and from the relative sediment loads obtained from the current measurement campaign.

Table 7. Parameters used for sediment calibration, initial and final ranges

SWAT Code	Unit	Description	Initial range	Final value
SPCON	-	Parameter defining maximum amount of sediment that can be re-entrained during channel sediment routing.	0.0001 – 0.01	0.0025
SPEXP	-	Exponent parameter for calculating sediment re-entrained in channel sediment routing	1.0 – 2.0	1.5
USLE_K	0.013*ton*m2*hr)/ (m3* ton* cm)	USLE equation soil erodibility (K) factor	0.02 – 0.30	0.04 – 0.25

The sediment inflow as simulated was compared with the sediment inflow that was derived from the bathymetric surveys and the calculations on the reservoir trap efficiencies and sediment consolidation (Table 7).

Table 8. Reservoir sediment inflow: simulated and calculated from bathymetric surveys

Reservoir	Sediment inflow - simulated Mtons/yr	Sediment inflow - calculated Mtons/yr	Simulated as % of observed %
Sasumua	0.0	0.0	N/A
Ndakaini	0.0	0.0	N/A
Masinga	6.3	8.0	80%
Kamburu	0.9	1.1	82%
Total	7.2	9.2	80%

* from Thiba catchment

It has to be noted for the Kamburu reservoir, the table below shows only the sediment contribution of the Thiba catchment and does not include sediment released by the Masinga reservoir. Especially during high flow periods, the Masinga reservoir may show a brownish colour, indicating the suspension of sediments, but due to the dimensions of the Masinga reservoir, only very fine particles are released. These fine particles do not have enough time to settle in the Masinga reservoir in spite of its large dimenstions. It can therefore be assumed that neither a significant fraction of it will settle in the Kamburu reservoir. Calculations done on the trap efficiency of both reservoirs confirm this assumption

As can be seen in the previous table, the sediment inflow simulated by the calibrated model is around 80% of the sediment inflow that is derived from the bathymetric surveys. This means that the vast majority of all the sediment inflowing into the reservoirs can be attributed to erosion from natural and cultivated land covers. The remainder is likely to be caused mainly by infrastructural works located upstream of the reservoirs (roads, culverts, construction projects, badly protected urban water and traffic infrastructure, etc.). These sediment sources are normally rather local and often also temporal.

Figure 24 shows for each of the monitored locations, the average observed sediment discharge during the current measurement campaign versus the average long-term sediment loads that were simulated. As can be seen, there is a clear correspondence between simulated and currently observed (R^2 = 95%). The good correspondence confirms that the model represents fairly well the relative contribution of each of the sub-catchments upstream of the monitored locations. On the other hand, the figure also strikes the highly dynamic behavior of the sediment response of the catchment. In spite of the fact that the current measurement campaign takes place during a relatively wet year, average sediment discharge is significantly higher, as is also confirmed by the bathymetric surveys.

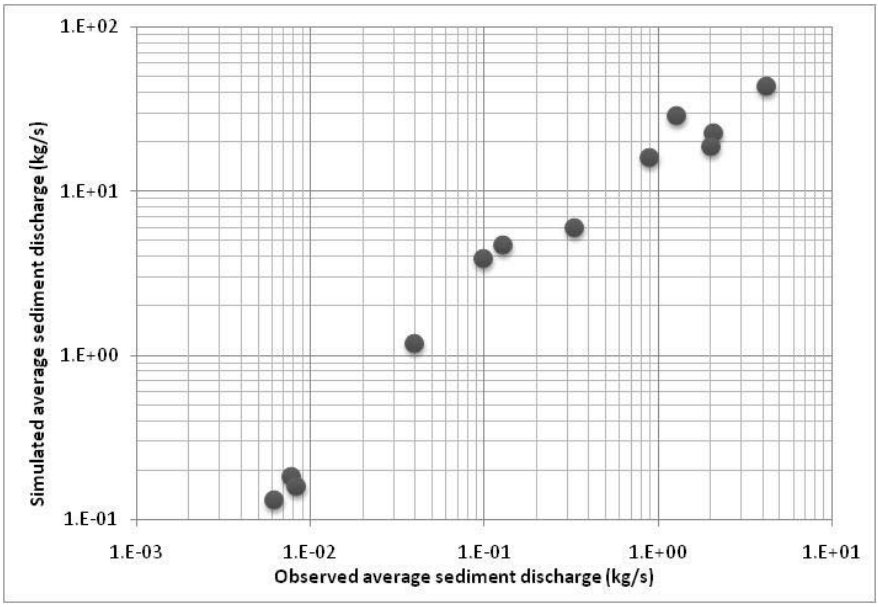

Figure 24. Average observed sediment loads during current measurement campaign versus average long-term sediment loads

The key objective of the modeling assessment is to support decision to be made on basin scale on future interventions. For this reason, it is essential that the model reproduces well the historical status of the basin, both in time as in space. The comparison with measured time series at different locations showed that the model represents fairly well the basin responses in terms of streamflows and sediments, given the scope and scale of the study. The good performance enables the analysis of future scenarios using the model, in order to plan future interventions and coordinate implementation efforts with the different actors in the basin.

4 Simulated baseline sediment budgets

The results from the modeling assessment, describing the current baseline situation is presented in the following sections. Project Outcome Indicators (POIs) and a temporal and spatial analysis give quantitative insight in the current state of the basin in terms of erosion and reservoir sedimentation. The POIs were also used to assess the impact of the different management and adoption scenarios in the following chapter. Besides, the section on targeting of the WRUAs includes a detailed table where the erosion rate is quantified for each management unit.

4.1 Basin-wide assessment

Two indicators are critical when characterizing the sediment dynamics in the basin, which are the erosion rate on the fields, expressed in the amount of sediment lost per unit area and unit time (f.e. ton/ha/yr), and the sediment yields in the stream network, expressed in Mton/yr. These two POIs can be defined using the model output on each location throughout the basin. This section provides the outcomes of these POIs for the main crops in the basin, and the two main reservoirs. The main crops in the basin are (i) coffee, (ii) subsistence crops, mainly maize, and in table denoted as maize; (iii) tea. Also the erosion rate is given for the entire basin, including the non-cultivated areas. It has to be noted that the other reservoirs in the basin are not included as the bathymetric survey indicated that no siltation takes place, as was also confirmed by the modeling exercise.

Table 10 shows the outcomes for the POIs, as extracted and post-processed from the 30-year output of the SWAT model. To give an indication of the temporal variability of these POIs, besides the 30-year average, also the standard deviation, the 10% percentile and 90% percentile are given, all based on yearly aggregated values.

Table 9. Project outcome indicators based on 30-year simulated values for the Upper Tana basin

	Average	Standard deviation	Percentile 10%	Percentile 90%
Erosion rate (ton/ha/yr)				
Coffee	27	15	6	46
Maize	18	9	4	28
Tea	14	6	5	20
Entire U. Tana	10	5	2	16
Sediment yields (Mton/yr)				
Inflow Masinga	6.3	3.7	1.1	10.6
Inflow Kamburu	0.9	0.6	0.4	1.6

As the table demonstrates, the yearly variability is large due to the rainfall regime in the basin. The temporal variability of the sediment reservoir inflow is better represented in Figure 25, which shows total basin yearly precipitation and the yearly reservoir volumes that enter Masinga and Kamburu reservoir. Total sediment yields entering Kamburu and Masinga range from less than 1 MCM during a dry year to more than 15 MCM during the wetter years.

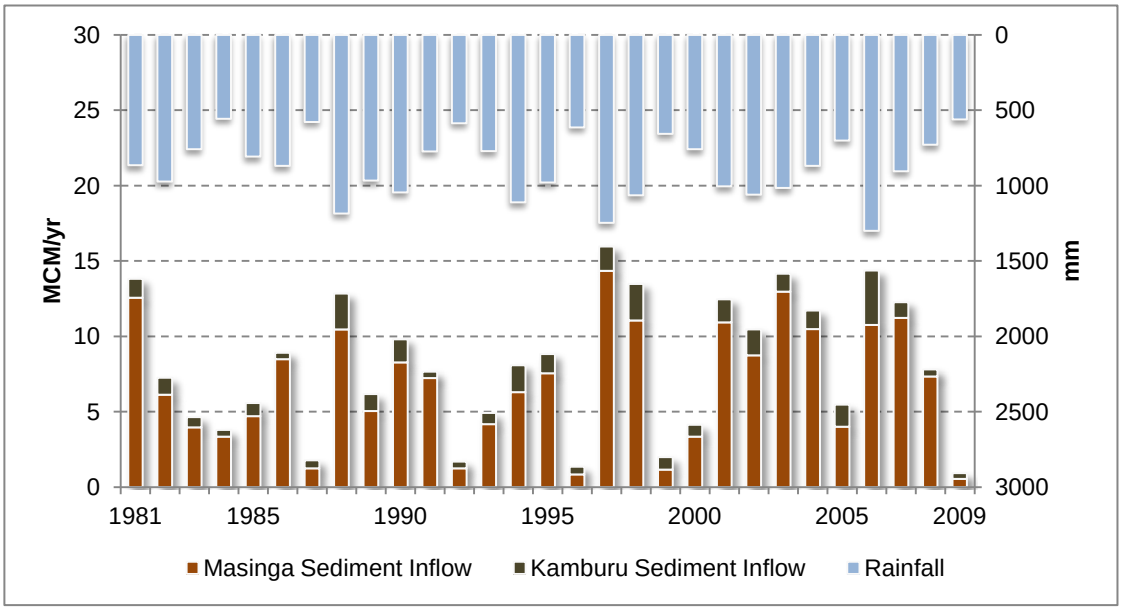

Figure 25. Yearly sediment inflow amounts for Kamburu (Thiba) and Masinga reservoir (Upper Tana and Thika) with basin average yearly rainfall.

This figure demonstrates that in spite of the clear relationship between annual rainfall and sediment flows, the highest sediment yields do not exactly correspond with the highest annual rainfall amounts. This is because actual erosion depends very much on the soil moisture state and previous weather conditions (compare 2003 and 2006). Another observation from this figure is that total sediment inflow can be more or less similar while the distribution (Kamburu versus Masinga) may be quite different (also compare for example 2003 and 2006).

4.2 Spatial analysis

The model is used to generate maps of erosion taking place throughout the basin. The map below shows the spatial distribution of the yearly gross erosion rate, averaged over 30 years (1980-2010). The erosion rates were classified according to the following scheme (similar to Singh et al, 1992):

- slight (0–2 t/ha/ yr),
- moderate (2–5 t/ha/ yr),
- high (5–10 t/ha/ yr),
- very high (10– 20 t/ha/ yr),
- severe (20–50 t/ha/ yr) and
- very severe (>50 t/ha/ yr)

As can be seen, large differences exist in terms of absolute erosion rates over space. Natural areas hardly contribute any sediment while other areas contribute the vast majority of the total sediment budget.

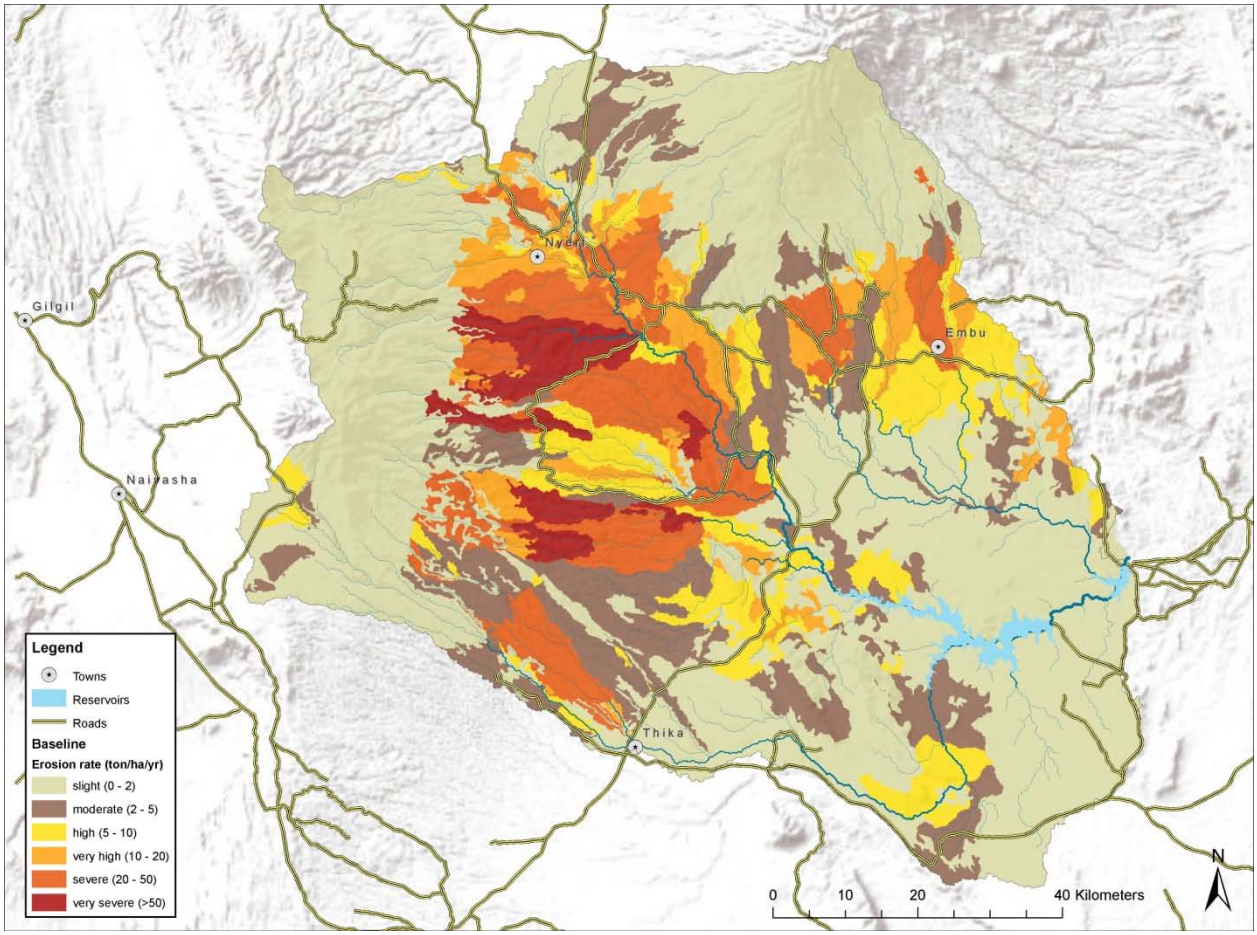

Figure 26. Spatial distribution of the average yearly gross erosion rate of the basin (see also Appendix)

The eroded materials produced in each part of the basin are routed through the channel network. The amount of sediments that reach downstream points of the basin depends on the channel deposition and degredation processes. These processes are simulated in the model which is able to provide an estimate of the sediment load of each reach segment.

Figure 27 shows the map (see also appendix) with the results from the model for all confluences in the basin of the average annual sediment discharge. This average sediment discharge value is calculated as the product of average yearly streamflow and average yearly sediment load, at each represented point. The color and dimension of the dots are proportional to the average sediment discharge simulated at that point. The influence of the reservoirs can be clearly observed for example at the Masinga reservoir where the dots become smaller just after the dam. The different ranges found in the basin are classified as follows:

- very low (< 0.2 Mton/yr)
- low (0.2-0.5 Mton/yr),
- medium (0.5-1.0 Mton/yr),
- high (1.0-2.0 Mton/yr),
- very high (> 2.0 Mton/yr)

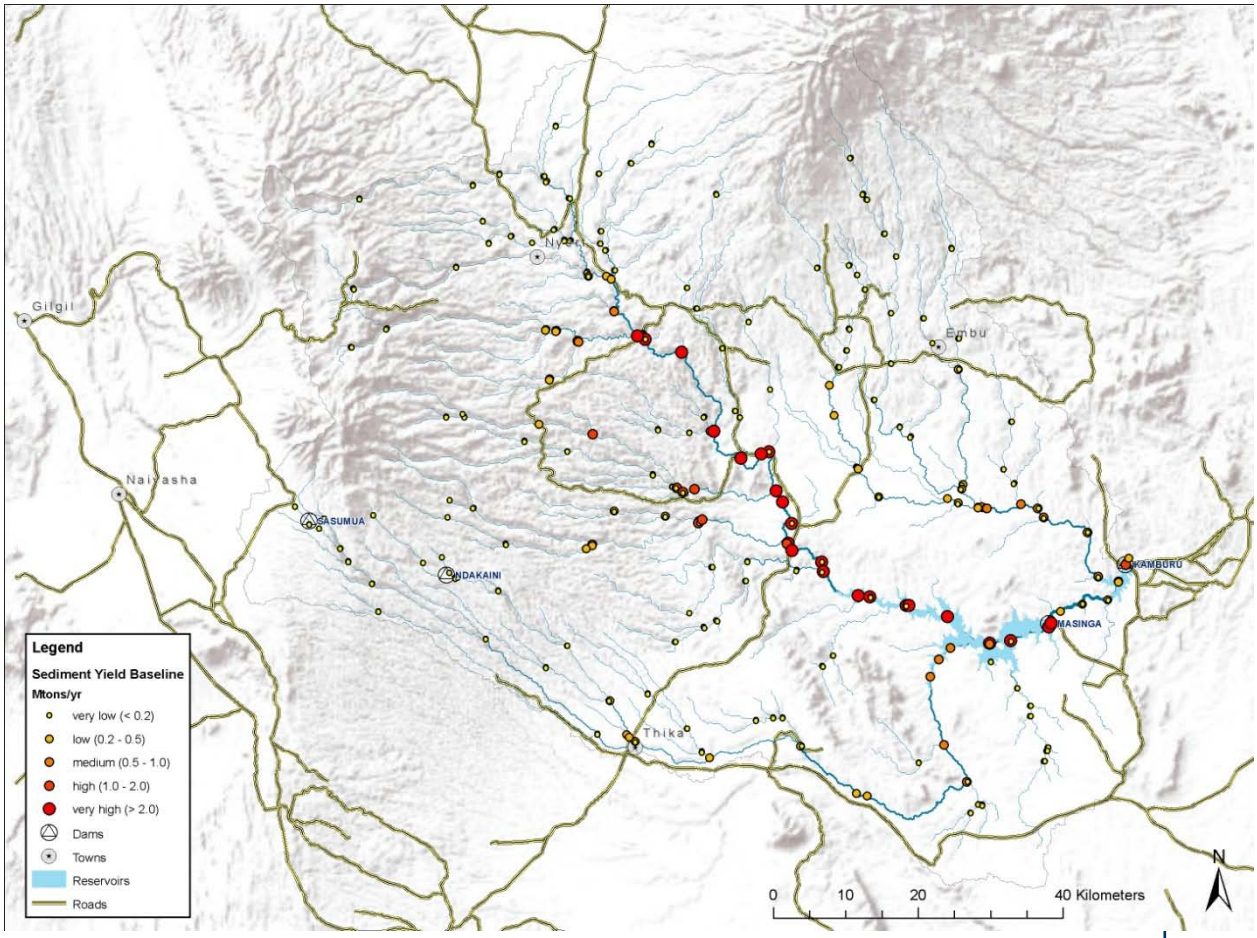

Figure 27. Average yearly sediment discharge (Mtons) at each confluence of the basin.

5 Simulated future sediment budgets

5.1 Description of studied interventions

WRMA aims at increasing the life span of the main reservoirs in the Upper-Tana basin by reducing sediment input into the main reservoirs. Several techniques and interventions allow reducing the erosion and sediment yields in the basin that can be studied using the erosion model. This chapter presents the results of the various intervention scenarios that were selected for this project during the inception phase. There principle objective is reducing erosion and sediment reaching the reservoirs.

The following techniques were studied within the future scenario analysis:

1. The use of vegetative contour strips, consisting of grass or other permanent vegetation in a contoured field to help trap sediment and nutrients. Because the buffer strips are established following the contours, runoff flows slower and evenly across the grass strip, reducing sheet and rill erosion. Permanent vegetative contour strips are in fact an inexpensive substitute for terraces.
2. Applying mulching, requiring residues produced within the cropping area and/or residues collected from elsewhere. These residues are applied in the field by spreading them on top of the soil. They protect to a certain extent the soil from erosion and reduce compaction from the impact of heavy rains.
3. Another technique consists of making soil ridges of varying width and height, average being 30cm width and 20 cm height. At regular intervals, crossties are built between the ridges. The ties are about two-thirds the height of the ridges, so that if overflowing occurs, it will be along the furrow and not down the slope. This technique is applicable in this area in the areas where subsistence crops are cultivated, being mainly maize.
4. Construction of check dams. These are small normally hand-made dams constructed in concentrated flow areas, not in streams. They reduce flow velocities and prevent gully erosion. Some removal of coarse to medium-size sediment is normally obtained, however, fine and clay particles pass. Inspection after storms is required, besides periodic repair and sediment removal. Gully restoration is recommendable to complement the check dam construction and ensure its effectiveness.

The first three interventions were studied as separate scenarios, but each of them combined with the fourth (check dams) technique. The first three require an effort of the farmer itself, while for the construction of check dams normally a community effort is required involving all the farmers cultivating their lands around a certain concentrated flow area or gully.

The following table summarizes the parameter changes that were incorporated into the model for each of the scenarios. For the construction of check dams, it was assumed that by implementing them into the concentrated flow areas, the effective slope steepness could be reduced by 5%.

Table 10. Parameter changes for a scenario with permanent vegetative contour strips

SWAT Parameter	Description	Land use	Baseline	Contour strips	Mulching	Tied ridges	Source
CN2	Runoff curve	Coffee	77	65	77	77	USDA-SCS

	number	Maize	77	70	77	62	(1986)
		Tea	70	65	70	77	
P_{USLE}	Support practice factor for soil loss	Coffee	1.0	0.7	0.8	1.0	Wischmeier and Smith (1978)
		Maize	1.0	0.9	0.8	1.0	
		Tea	1.0	0.7	0.8	1.0	
ESCO	Soil evaporation compensation coefficient	Coffee	0.95	0.95	0.99	0.95	Kannan et al. 2007
		Maize	0.95	0.95	0.99	0.95	
		Tea	0.95	0.95	0.99	0.95	

5.2 Comparison of interventions

The scenario changes were incorporated into the model and simulations were run over the period 1980-2010. Consequently, the scenarios were compared among each other and contrasted against the baseline scenario. Each of the scenarios include besides the famers' management practices, the construction of the low-cost infrastructure in the concentrated flow areas (check dams) that cause reduced flow velocities and inhibit to a certain extent the gully and bank erosion.

This part of the scenario analysis assumes that the management practices are implemented at all the fields where coffee, maize and tea are currently cultivated and check dams are constructed in all the concentrated flow areas of these fields. Based on the output of these scenarios, the WRUAs will be targeted for implementation and different adoption scenarios are run, as shown in the following sections. The POIs of these scenarios are shown in Table 10 allowing a comparison with the baseline scenario and among each other. Relative percentual differences compared to the baseline scenario are shown between brackets. A graphical representation of these outcomes is shown in Figure 28.

Table 11. Project outcome indicators demonstrating basin-wide impact of studied management scenarios

	Baseline	Contour strips	Mulching	Tied ridges	Check dams only
Erosion rate (ton/ha/yr)					
Coffee	27	9	21	27	
Maize	18	10	15	7	
Tea	14	6	10	14	
Entire Upper-Tana	10	5	8	7	
Sediment yields (Mton/yr)					
Inflow Masinga	6.3	3.4 *(-46%)*	5.6 *(-10%)*	4.6 *(-26%)*	*(-3%)*
Inflow Kamburu	0.9	0.4 *(-57%)*	0.7 *(-20%)*	0.6 *(-32%)*	*(-7%)*

Basin-wide implementation of vegetative contour strips produces the highest sediment yield reductions compared to the other scenarios. The scenario included the construction of check dams, of which the relative contribution is around 5%, as shown in the last column of the table. Basin-wide use of mulching reduces the sediment reservoir inflows by 10% to 20%, which is the lowest rate compared to the other scenarios. The implementation of tied ridges in the fields cultivated with subsistence crops (maize) can reduce the sediment inflow significantly. The reduction in this case is only obtained by intervening in the subsistence cropped areas, and not in the coffee and tea zones, as described in the previous section.

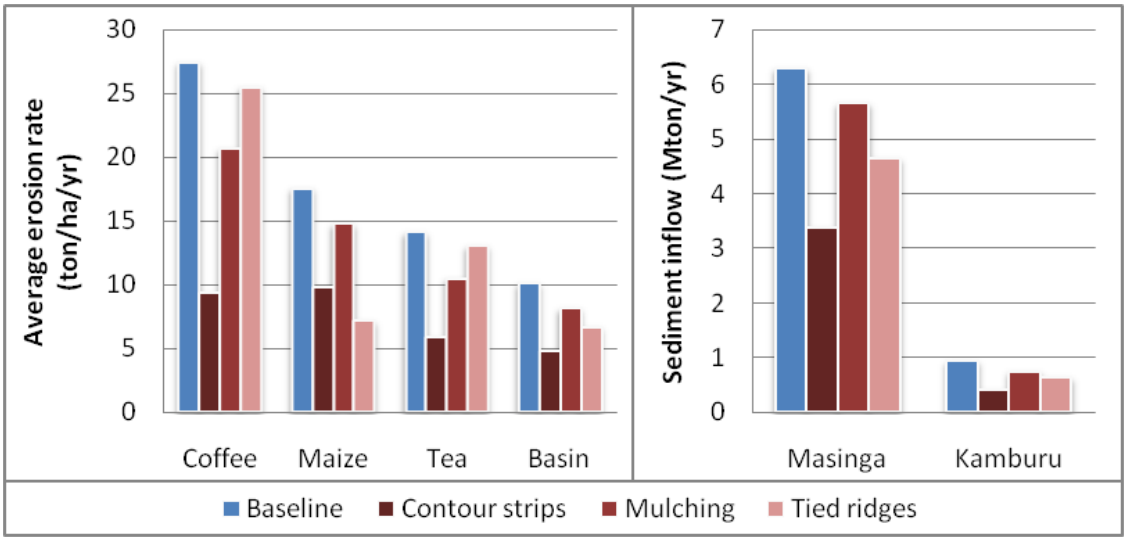

Figure 28. (a) Impact on erosion rate per land use; (b) Impact on sediment inflow into main reservoirs

The model provides information on the erosion and sediment response for each of the spatial calculation units throughout the basin, being more than 2000. The results therefore provide detailed spatial insight in the areas where most erosion takes place. The same classification as for the baseline situation in Figure 26 was used to map the erosion distribution of the three scenarios. Figure 29 gives a snapshot of the vegetative contour strips scenario (right) compared to the baseline scenario (left). For more detailed maps, see the Appendices.

As can be seen clearly, the reddish colours indicating severe erosion in the area are less apparent in the scenario run. This confirms that the practices have the potential to reduce the gross erosion within the basin significantly in the erosion-prone areas. These areas are located mostly in the coffee and maize zones on the steep slopes of the Aberdares, and to a less extent on the Mount Kenya slopes. The following section uses this spatial information to target the existing water user associations (WRUAs) for implementation.

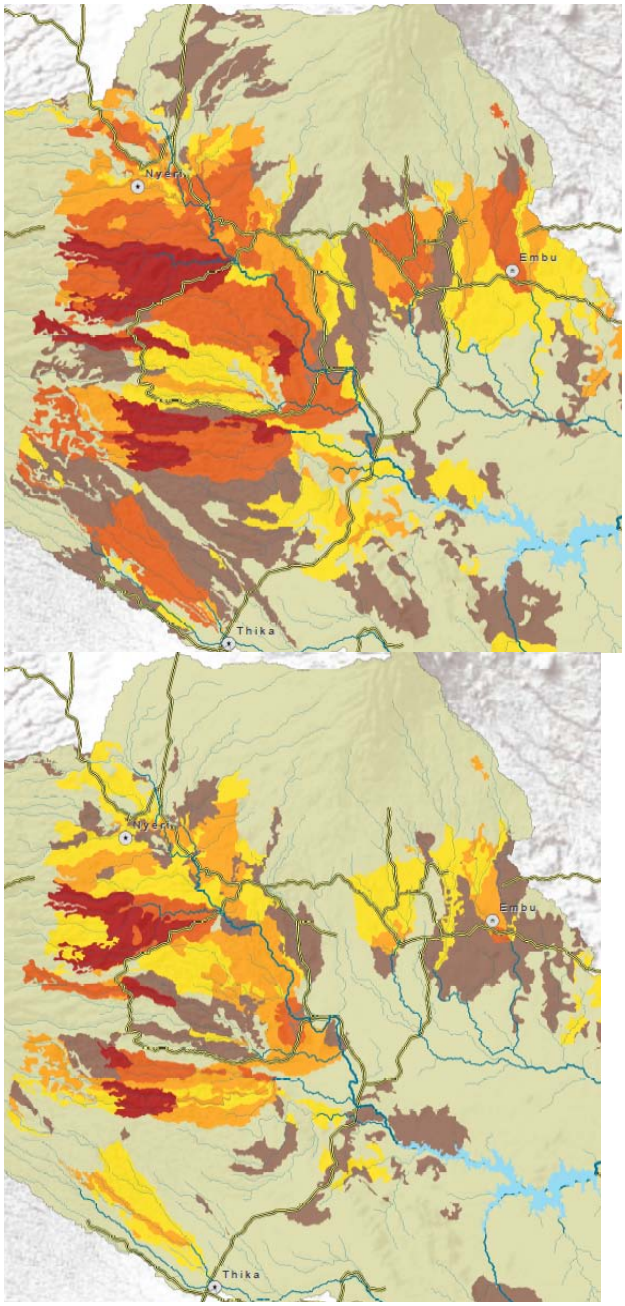

Figure 29. Difference in average gross erosion rate (ton/ha/yr) between the baseline and the scenario with permanent vergetative contour strips (see Appendices for detailed maps)

5.3 Targeting of the WRUAs for implementation

It is agreed that WRMA will, both at the regional and sub-regional level, involve the WRUAs in meeting the objective in the following years to reduce erosion and sediment yields. Also selection and marking the exact sites for the interventions requires local information of the WRUAs where possible. Besides, WRUAs are the organisms in which farmers are best organized in this particular basin, promising the highest level of adaptation of the proposed interventions. Nowadays, 25 WRUAs are currently operational in the Upper-Tana basin (see Figure 30) and they cover around 50% of the total cultivated area.

Targeting and prioritizing the WRUAs for implementation of the practices ensures that the objective of reducing sediment yields in the basin can be obtained in the most cost-effective way. The spatial model output is used to carry out this assessment.

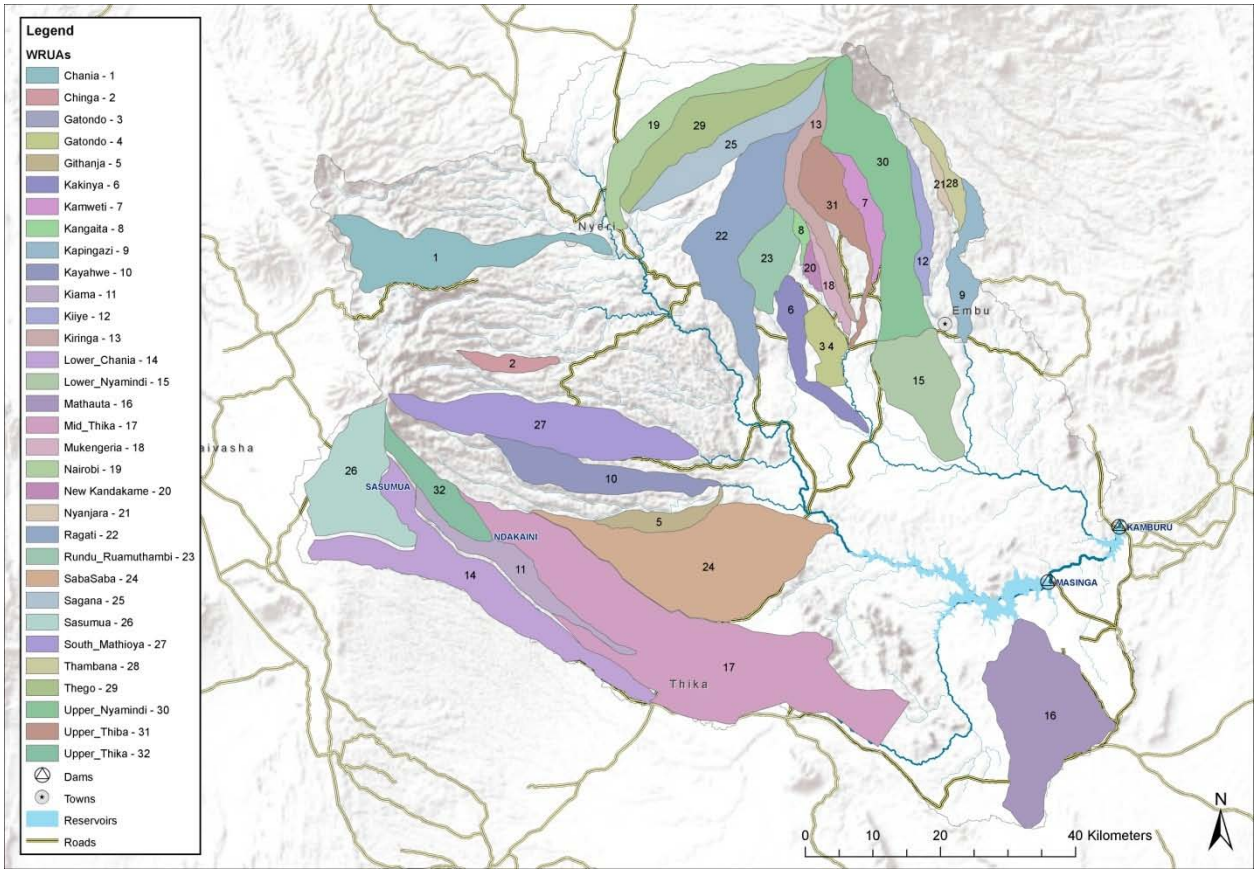

Figure 30. Map of the WRUAs in the Upper-Tana basin. The numbers correspond to the legend entries

The prioritization of the WRUAs depends on the scenario studied. In order to integrate all scenarios, the maximum erosion reduction was taken for each calculation unit of the model. The reductions of all the calculation units belonging to the territory of each WRUA were averaged, as shown in Table 11. Then they were ranked based on these averages and given a priority level according to their ranking score. The priority levels were assigned according to the following scheme:

- Level I: reductions in erosion rate of more than 10 ton/ha/yr
- Level II: reductions in erosion rate of 7-10 ton/ha/yr
- Level III: reductions in erosion rate of 5-7 ton/ha/yr
- Level IV: reductions in erosion rate of less than 5 ton/ha/yr

Table 12. Maximum erosion reductions (ton/ha/yr) obtained by the scenarios for each WRUA and relevant land use.

WRUA	Coffee	Maize	Tea	Average	Area (km2)	Priority
Chania	-6	-9		-7	43	II
Chinga		-39	-12	-32	21	I
Gatondo		-3		-3	12	IV
Kakinya	-7	-1		-3	70	IV

Kapingazi	-9	-9		-9	40	II
Kayahwe	-36	-38	-7	-31	73	I
Kiama		-23		-23	5	I
Lower Chania	-2	-2	0	-1	93	IV
Lower Nyamindi		-5		-5	71	III
Mathauta		-1		-1	194	IV
Mid Thika	-9	-3	-3	-7	418	II
Mukengeria			-3	-3	3	IV
Nairobi		-4		-4	14	IV
New Kandakame	-16		-1	-8	40	II
Ragati	-8	-8	-1	-6	100	III
Rundu Ruamuthambi	-4		-1	-2	107	IV
SabaSaba	-4	-5	0	-5	466	III
Sagana	-7	-13	-1	-7	33	II
Sasumua		-3		-3	80	IV
South Mathioya	-18	-19	-2	-12	231	I
Thambana			-2	-2	28	IV
Thego	-6	-2		-3	70	IV
Upper Nyamindi	-9	-8	-1	-8	148	II
Upper Thiba	-12	-13	-2	-10	48	I
Upper Thika			-6	-6	59	III

***Cultivated area**

5.4 Implementation and adoption scenarios

Consequently, scenarios were run assuming different adoption levels of the scenario that showed the highest basin-wide reduction in sediment inflow. By studying different levels of adoption, WRMA will be able to decide in the following-up of this project, which effort and budget will be needed depending on the objectives and predicted outcomes.

As shown before in the comparison of the interventions, the scenario with the highest basin-wide reduction in sediment inflow was the one that incorporated the use of vegetative contour strips by the farmers, mainly in the coffee and maize fields, combined with the restoration of gullies, including check dams in the concentrated flow areas.

The adoption scenarios were run with three different levels, using the priority levels defined before when targeting the WRUAs:

- Adoption level 1: Adoption of the practices and implementation of constructive and restorative works by the WRUAs given priority **I**, corresponding to around **8%** of the total cultivated area of the basin and involvement of 5 WRUAs
- Adoption level 2: Adoption of the practices and implementation of constructive and restorative works by the WRUAs given priority **I and II**, corresponding to around **20%** of the cultivated area of the basin and involvement of 11 WRUAs
- Adoption level 3: Adoption of the practices and implementation of constructive and restorative works by the WRUAs given priority **I, II and III**, corresponding to around **35%** of the cultivated area of the basin and involvement of 15 WRUAs

Table 12 and Figure 31 show the results of the three adoption scenarios.

Table 13. Project outcome indicators according to adoption level scenarios

	Baseline	Adoption level 1	Adoption level 2	Adoption level 3
Erosion rate (ton/ha/yr)				
Coffee	27	25	23	22
Maize	18	17	16	16
Tea	14	13	13	12
Entire Upper-Tana	10.1	9.5	9.1	8.8
Sediment yields (Mton/yr)				
Inflow Masinga	6.3	5.8	5.6	5.3
Inflow Kamburu	0.9	0.9	0.7	0.7

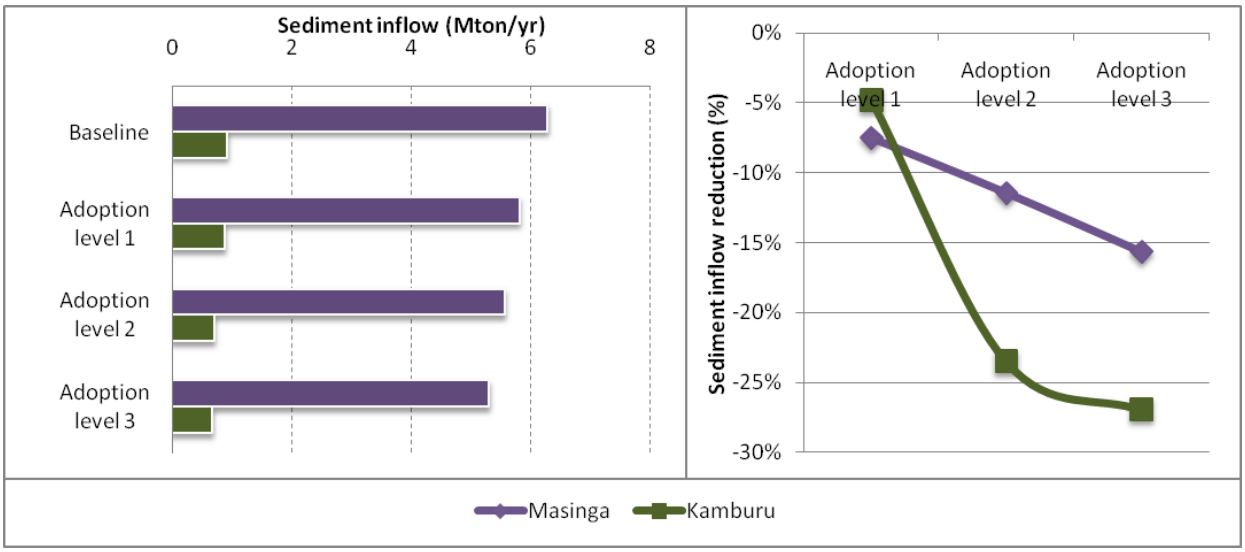

Figure 31. Sediment inflow and relative reductions according to the level of adoption and implementation

The following conclusions can be drawn, based on the outcomes of the scenarios in which different levels of adoption were assessed:

i. By implementing vegetative contour strips in the areas that belong to the WRUAs, given priority I for implementation, a sediment reservoir inflow reduction can be obtained of around 5%.

ii. Adoption levels 2 and 3, assuming implementation in all the cultivated areas that belong to the WRUAs given priority I, II and III, would reduce sediment input into the Masinga reservoir by about 15%, while the Kamburu reservoir sediment inflow would be reduced by around 25%.

iii. A successful involvement of the WRUAs to support the interventions to be taken, can play a significant role in reducing sediment yields and the sedimentation of the reservoirs.

6 Concluding remarks on model outcomes

The following conclusions can be drawn specifically on the outcomes of the baseline and scenario analysis using the model that was setup using state-of-the-art datasets and calibrated with historic data and data from the monitoring activities of this project.

1. The model that was set up and calibrated for this project was able to reproduce well the sedimentation of the reservoirs that occurred since construction. The bathymetric surveys gave a precise figure on the total capacity loss that could be enhanced by the model providing detailed spatial and temporal output.
2. Annual sediment flows relate well with annual rainfall although generally the sediment response has a higher variability than the rainfall itself. Erosion rates are highest in the coffee zones and areas where maize and other cereals are cultivated. Intervening in particularly these areas is therefore considered optimal.
3. The scenario analysis showed that the highest basin-wide reduction on reservoir sediment inflow can be expected from the vegetative contour strips scenario. Especially implementation of these practices together with the check dams at the coffee fields has the potential to reduce erosion dramatically.
4. Mulching has only little effect on the sediment budgets in the basin, as runoff is hardly reduced by this practice. Most of the impact of this scenario can be attributed to the effect of the check dams which reduce velocities in the concentrated flow areas.
5. Tied ridges are not feasible for coffee and tea zones, so this practice was only studied for the fields where subsistence crops, maize and other cereals are cultivated. On these fields, erosion reductions can be obtained that are even larger than when vegetative contour strips are implemented. Maintenance of this practice however is more time-consuming and costly than the contour strips. It should however by considered when starting the coordinative activities following this project with the WRUAs.
6. Gully restoration and the construction of check dams to complement the farmers' practices was studied in all the scenarios and is highly recommended to reduce peak flows and degradation of gullies and channels.
7. Based on the scenario analysis, priority levels were given to the areas that correspond to the currently existing WRUAs in the basin. The 25 WRUAs cover around 50% of the total cultivated area in the Upper Tana. Different adoption levels of the most effective scenario was studied to support WRMA in the following up phase for implementation by involving the WRUAs. The analysis showed that the involvement of around 5 WRUAs would reduce sediment reservoir inflow by around 5-8%, while implementation at the territories of 15 WRUAs could reduce reservoir sedimentation by around 15-25%, increasing the life expectancy of the reservoirs accordingly.

7 References

Asadullah, A., N. McIntyre, M. Kigobe, 2008. Evaluation of five satellite products for estimation of rainfall over Uganda. Hydrological Sciences Journal 53: 6.

Batjes NH and Gicheru P 2004. Soil data derived from SOTER for studies of carbon stocks and change in Kenya (ver. 1.0). Report 2004/01, ISRIC - World Soil Information, Wageningen

Brune G.M. (1953) Trap efficiency of reservoirs. Transactions-American Geophysical Union, 34(3), 407–418.

Churchill M.A. (1948) Analysis and use of reservoir sedimentation data. Proceedings of the Federal Interagency Sedimentation Conference, US Bureau of Reclamation: Denver.

Droogers, P., and G. Kite. 2001. Simulation modeling at different scales to evaluate the productivity of water. Physics and Chemistry of the Earth 26: 877-880.

Droogers, P., M. Torabi, M. Akbari, and E. Pazira. 2001. Field scale modeling to explore salinity in irrigated agriculture. Irrigation and Drainage 50: 77-90.

Funk C, Verdin J (2003) Comparing satellite rainfall estimates and reanalysis precipitation fields with station data for western Kenya. Proceedings of the International Workshop on Crop Monitoring for Food Security in Africa, European Joint Research Centre/UN Food and Agriculture Organization, Nairobi, Kenya, January 28–30

Funk C, Verdin J (2010) Real-Time Decision Support Systems: The Famine Early Warning System Network. In: Satellite Rainfall Applications for Surface Hydrology, Gebremichael, M. and Hossain, F. (Eds.), DOI 10.1007/978-90-481-2915-7 Publ. Springer

Hunink, J.E., W.W. Immerzeel, P. Droogers, 2009. Green Water Credits for the Upper Tana Basin, Kenya. Phase II - Pilot Operations: Biophysical assessment using SWAT. Report FutureWater: 84

JICA, 1992. The Study on the National Water Master Plan. Prepared by the Ministry of Water Development with the assistance of Japan International Cooperation Agency (JICA).

KSS 1996. The Soil and Terrain Database for Kenya at scale 1:1,000,000 (ver. 1.0), Kenya Soil Survey, National Agricultural Laboratory, Kenya Agricultural Research Institute

KSS and ISRIC 2007. Kenya Soil and Terrain database - version 2. Kenya Soil Survey and ISRIC - World Soil Information

NASA 1998. The NASA GSFC and NIMA (National Imagery and Mapping Agency) Joint Geopotential Model EGM96 (http://cddis.nasa.gov/926/egm96/egm96.html)

Parajuli, P.B., K. R. Mankin, and P.L. Barnes. 2008. Applicability of targeting vegetative filter strips to abate fecal bacteria and sediment yield using SWAT. Agric. Water Manage. 95: 1189-1200.

Singh G; Ram Babu; Pratap Narain; Bhushan L S; Abrol I P, 1992. Soil erosion rate in India. Journal of Soil and Water Conservation, 47(1), 97–99

Sobierja JA, Elsenbeer H and Vertessy RA 2001. Pedotransfer functions for estimating saturated hydraulic conductivity: implications for modeling storm flow generation. Journal of Hydrology 251, 202-220

Sombroek WG, Braun HMH and van der Pouw BJA 1982. Exploratory soil map and agro-climatic zone map of Kenya, 1980, scale 1:1,000,000. Exploratory Soil Survey report No E1. Kenya Soil Survey, Ministry of Agriculture, Nairobi

Tripathi, M.P., R.K. Panda, and N.S. Raghuwanshi. 2003. Identification and Prioritisation of Critical Sub-watersheds for Soil Conservation Management using the SWAT Model. Biosystems Engineering (2003) 85(3):365-379, doi:10.1016/S1537-5110(03)00066-7.

van Engelen VWP and Wen TT 1995. Global and National Soil and Terrain Digital Database (SOTER). Procedures Manual. ISRIC - World Soil Information, Wageningen

Van Engelen VWP, Batjes NH, Dijkshoorn K and Huting J 2005. Harmonized Global Soil Resources Database (Final Report). Report 2005/06, ISRIC - World Soil Information, Wageningen

Wooldridge.R. (1984) Sedimentation in reservoirs: Tana River basin, Kenya. Ill -Analysis of hydrographicsurveys of three reservoirs in June/July 1983. Report no. 61, Hydraulics Research, Wallingford, UK

WRMA 2010. Consultancy Services On Micro Projects Investments In The Upper Tana Catchment. Stage 11. Mt. Kenya Sub-Catchment Management Plan, March 2010.

Xie P, Arkin PA (1997) A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bulletin of the American Meteorological Society 78(11):2539–2558

Yang, J., Abbaspour K. C., Reichert P., and Yang H. 2008. Comparing uncertainty analysis techniques for a SWAT application to Chaohe Basin in China. In review. Journal of Hydrology. 358(1-2):1-23.