Hydrological Modeling in Data-Scarce Catchments: The Kilombero Floodplain in Tanzania

Kristian Näschen 1,*, Bernd Diekkrüger 1, Constanze Leemhuis 1, Stefanie Steinbach 1, Larisa S. Seregina 2,3, Frank Thonfeld 1 and Roderick van der Linden 2,3

1 Department of Geography, University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany;
   b.diekkrueger@uni-bonn.de (B.D.); leemhuis@uni-bonn.de (C.L.); s.steinbach@uni-bonn.de (S.S.); frank.thonfeld@uni-bonn.de (F.T.)
2 Institute for Geophysics and Meteorology, University of Cologne, Pohligstr. 3, 50969 Cologne, Germany;
   lseregin@uni-koeln.de (L.S.S.); rvdlinde@uni-koeln.de (R.v.d.L.)
3 Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany

* Correspondence: knaesche@uni-bonn.de; Tel.: +49-228-73-3335

Received: 3 March 2018; Accepted: 1 May 2018; Published: 4 May 2018

Abstract: Deterioration of upland soils, demographic growth, and climate change all lead to an increased utilization of wetlands in East Africa. This considerable pressure on wetland resources results in trade-offs between those resources and their related ecosystem services. Furthermore, relationships between catchment attributes and available wetland water resources are one of the key drivers that might lead to wetland degradation. To investigate the impacts of these developments on catchment-wetland water resources, the Soil and Water Assessment Tool (SWAT) was applied to the Kilombero Catchment in Tanzania, which is like many other East African catchments, as it is characterized by overall data scarcity. Due to the lack of recent discharge data, the model was calibrated for the period from 1958–1965 ($R^2 = 0.86$, $NSE = 0.85$, $KGE = 0.93$) and validated from 1966–1970 ($R^2 = 0.80$, $NSE = 0.80$, $KGE = 0.89$) with the sequential uncertainty fitting algorithm (SUFI-2) on a daily resolution. Results show the dependency of the wetland on baseflow contribution from the enclosing catchment, especially in dry season. Main contributions with regard to overall water yield arise from the northern mountains and the southeastern highlands, which are characterized by steep slopes and a high share of forest and savanna vegetation, respectively. Simulations of land use change effects, generated with Landsat images from the 1970s up to 2014, show severe shifts in the water balance components on the subcatchment scale due to anthropogenic activities. Sustainable management of the investigated catchment should therefore account for the catchment–wetland interaction concerning water resources, with a special emphasis on groundwater fluxes to ensure future food production as well as the preservation of the wetland ecosystem.

Keywords: SWAT model; hydrological modeling; East Africa; land use changes; water balance; wetlands

1. Introduction

Recent developments show an increasing trend to utilize wetlands in East Africa [1–4]. This development is triggered by several issues, such as increasing food demand caused by demographic growth, climate change, and degradation of upland soils. Unlike the uplands, wetlands hold potential for year-round harvest, due to their fertile soils with a balanced soil water availability throughout the year [5].

In Tanzania the “Kilimo Kwanza” (Agriculture First) policy of the government prioritizes agricultural development and leads to agricultural intensification and expansion [6]. This increased utilization results in an intensive use of wetland resources, and may lead to a degradation of the
wetland system [7,8]. The understanding of the functioning of a wetland lies in the water budget, hence, information on water resources is essential to establish a sustainable land management system within a wetland. Without sufficient water, agricultural production deteriorates and food security is endangered [9]. For a holistic view on water resources and their fluxes, it is not sufficient to consider only processes within the wetland—one also has to consider the hydrological processes occurring within the entire catchment.

A typical approach to adequately calculate and represent hydrological processes within a catchment is the application of a hydrological model. Process-based models like the water balance simulation model (WaSiM) [10], SWAT [11], or MIKE SHE [12] are capable of simulating water resources under changing environmental conditions, such as climate change or land use and land cover changes (LULCC). A drawback of these models is the intensive data demand due to their physically-based approach. Unfortunately, data scarcity is an obstacle with regard to hydrological modeling in East Africa. Local water authorities face numerous challenges, like accessibility of discharge stations (especially in the rainy season), limited staff, and insufficient equipment due to restricted funds [6]. This implies heterogeneous datasets, with a low quantity and quality of the available data with regard to spatial and temporal coverage. Besides the hydrological data, other geodata, like topography, land use/land cover (LULC), soil, or climate data is needed to simulate biophysical processes. Local biophysical data (e.g., vegetation, soil, climate) is rarely available, but these gaps can be closed by applying remote sensing data, when operating on meso- to macroscale catchments [13,14].

Nevertheless, scientific guidance is particularly needed in data-scarce regions of East Africa, to assist with water resource management on the catchment scale [15]. In particular, the increased pressure on wetlands and their surroundings might alter the distribution and amount of water resources through LULCC. Hence, a spatially-distributed hydrological model is needed, to simulate the water balance. The capability of SWAT to simulate the water balance in data-scarce tropical regions of different scales has been proven in various studies [16–19], but only a few within East Africa [15,20–23].

This study was undertaken in the Kilombero Catchment in Tanzania. Although there is an urgent need to gather a detailed understanding of the water resources in the Kilombero Catchment, only a few attempts have been made to understand the hydrological system. Yawson et al. [24] applied lumped models and linear transfer approaches, while Lyon et al. [25] investigated spatio-temporal drainage patterns. Subcatchments of the Kilombero Catchment were analyzed by Burghof et al. [26], Koutsouris et al. [27], and Daniel et al. [28], by developing a conceptual model, setting up an HBV (Hydrologiska Byråns Vattenavdelning) light model [29] for two subcatchments and by analyzing soil hydrological properties in a floodplain transect. These studies investigated water-related relationships within the Kilombero Catchment, giving either a detailed insight into specific subcatchments without considering the entire catchment [25–27]—or the entire catchment was analyzed, but without considering the spatial heterogeneity [24]. This spatial heterogeneity needs to be considered, according to Lyon et al. [25], who demonstrated the importance of the spatio-temporal variations in hydrological processes within the catchment. Nevertheless, Lyon et al. [25] applied recession analysis only as a kick-off for more detailed hydrological modeling. Altogether these studies did very good preparatory work, but there is still a research gap concerning the understanding of the distributed hydrological processes in the entire Kilombero Catchment. This research tries to bridge this gap by establishing the semi-distributed SWAT model, to gather a better spatial understanding of the hydrological processes in the Kilombero Catchment.

The main objective of this study is therefore to enhance the knowledge about the hydrological system of the Kilombero Catchment, in order to support water management in this changing environment [30,31]. Accordingly, three specific objectives were formulated:

(i) Assessing LULCC in the Kilombero Catchment since the 1970s;
(ii) Setting up a distributed hydrological model suitable to simulate impacts of LULCC;
(iii) Analyzing the impacts of LULCC on water balance components in the catchment.
These objectives are achieved by combining available local data with freely available global geodata, in order to adequately represent the hydrological system with the SWAT model. To obtain a better spatio-temporal understanding of the water balance components, the SWAT model is calibrated and validated against measured historical discharge data, and a spatio-temporally detailed water balance analysis on the catchment and subcatchment scale is presented. The observed impacts of LULCC from 1970 to 2014 were analyzed by deriving land use maps from Landsat images and implementing them into the calibrated model [30].

2. Materials and Methods

2.1. Study Site

The Kilombero Catchment is located in south-central Tanzania (Figure 1). The catchment is characterized by high relief energy, with altitudes ranging from 200 m to 2500 m above sea level, and is surrounded by the Udzungwa Mountains in the north, as well as the Mbarika Mountains and the Mahenge Highlands in the southeast (Figure 1). In total, the catchment comprises 40,240 km$^2$ up to the confluence of Kilombero and Rufiji River. Although the Kilombero Catchment only covers 23% of the drainage area of the Rufiji Basin, it contributes 62% of the annual runoff volume [32]. The floodplain system covers an area of 7967 km$^2$ [33], and contains the largest freshwater wetland within East Africa below a threshold of 300 m above mean sea level [34]. A big share of the floodplain is designated as a Ramsar site, which underlines the wetland’s international environmental importance [32,35].

The Kilombero River is the main tributary of the Rufiji River, representing the largest river basin in Tanzania. Water resources monitoring is scarce in the Kilombero Catchment, although it is prone to environmental changes with implications on water availability. Recent developments show an increase in population and agricultural land, and a decrease of natural landscapes, especially in the lower floodplain wetland of the catchment, while the upper catchment area is undergoing deforestation activities [30,32]. Up to today, 9% of the national rice yield is produced in the Kilombero wetland, and the wetland area is characterized by patches of several land use activities, from small- and large-scale farmers to pastoralists and urban populations near the town of Ifakara at the northeastern bottleneck of the wetland [36]. All these anthropogenic activities, in combination with ongoing climate change, alter the hydrological regime of the Kilombero River. Future activities foresee the establishment of an agricultural growth corridor in the Kilombero Catchment [37], which will foster the pressure on water resources in terms of quantity and quality in the research area and for downstream riparians.

The regional climate is defined as sub-humid tropical climate [38], with distinct dry and rainy (November–May) seasons with a predominantly unimodal rainfall pattern [34,38]. Nevertheless, many teleconnections influence the regional climate, resulting in shifts between unimodal and bimodal rainfall patterns among the years. Years with a unimodal distribution of rainfall lack the short rains (November–January), whereas the bimodal rainy seasons are characterized by short (November–January) and long rains (March–May) [32], which correspond mainly to the movement of the Intertropical Convergence Zone (ITCZ) [39]. The average annual precipitation is between 1200 and 1400 mm [38], with strong interannual variability [40] and spatial variability between the mountainous area and the lowlands, with precipitation up to 2100 mm and 1100 mm, respectively [32]. The temperature mirrors this pattern inversely, with annual mean temperatures of 24°C in the valley and 17°C in the uplands [32].

According to the Harmonized World Soil Database (HWSD) [41], the catchment is predominantly characterized by Fluvisols in the valley bottom, whereas the upland regions are dominated by Acrisols and Nitisols. The western upland soils are mainly described by Lixisols, and in the lower eastern part Cambisol is the dominant soil type (Figure 2).

The land cover of the upper catchment embraces a mixture of natural vegetation like tropical rainforests, bush lands, and wooded grasslands, with some patches of agricultural fields [42].
The valley is surrounded by a Miombo woodland belt, whereas the floodplain itself is dominated by agricultural use and grassland [43,44].

![Figure 1. Overview map of the study area, including available precipitation and discharge stations as well as the 0.44° Coordinated Regional Downscaling Experiment (CORDEX) Africa grid.](image)

2.2. Input Data

Data scarcity is a major problem in East Africa, and therefore a mixture of freely available global geo datasets and local measurements were combined and processed to run the hydrological model, following the approach of Leemhuis et al. [1]. Model calibration and validation is difficult because of low data availability, caused by the challenges of installing and maintaining the hydrometeorological network [6] for the Rufiji Basin Water Board (RBWB). The longest period with good quality discharge data was monitored from 1958–1970. The station utilized in this study is named Swero (Figure 1) and comprises 34,000 km$^2$, representing roughly 84% of the entire catchment area. The discharge time series of the Swero station ends in 1970, and therefore excludes the application of precipitation estimates from satellites. Hence, available station data from that period was included in the model, although the spatial distribution of precipitation stations is limited (Figure 1). Nevertheless, the temporal availability of the precipitation data for this study was good, with only 5.15% missing data for all stations and the entire simulation period. The other climate variables (Table 1) were taken from the
Temperature data was bias-corrected with the ERA-Interim reanalysis [46], by using the differences of the mean annual cycle of the 11-day running mean between each CORDEX model and the reanalysis. To account for the different topography, due to the different horizontal resolutions of the ERA-Interim and CORDEX models (namely 0.75° versus 0.44°), the correction was based on potential temperatures on the 700-hPa level. All climate variables, except precipitation, were taken as an ensemble mean of six historic regional climate model runs (Table 2). These six models represent a broad range of precipitation signals, with increasing, decreasing, and constant precipitation patterns when comparing the periods from 1986 to 2005 with 2040 to 2059.

Due to the lack of suitable Landsat images for one single year in the 1970s, a mosaic of eight Landsat pre-collection Level 1 images from the 1970s, downloaded from EarthExplorer [47], was classified using the supervised Random Forest classification [48] to adequately represent the land cover characteristics for the simulation. The 1994, 2004, and 2014 LULC maps are based on Landsat TM, ETM+, and OLI Surface Reflectance Level 2 Science Product imagery [49,50]. Due to recurrent extensive cloud cover, a year before and a year after the year in question were considered in addition. All scenes with <80% cloud coverage were cloud masked and additionally orthorectified where necessary, and the tasseled cap components wetness, greenness, and brightness [51–53], and the Normalized Difference Water Index (NDWI) [54] and Normalized Difference Vegetation Index (NDVI) [55] were calculated to account for seasonal dynamics of specific land cover classes. Based on the 30 m resolution Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) [56], we calculated slope and four morphometric indices (the terrain ruggedness index (TRI) [57], slope variability (SV) [58], topographic position index (TPI) [59,60], and the topographic wetness index.
The spectral multi-temporal metrics in combination with the DEM, slope, and topographic indices were classified using a random forest approach [48]. For the supervised classification, as well as for map validation, different reference datasets gathered on the ground in combination with high-resolution remotely-sensed data were used according to their availability. In order to complete the dataset, freely available data with varying spatial resolution were applied (Table 2). The changes in land cover from 1970 to 2014 are shown in Figure 3.

Table 1. Applied datasets, their resolution, sources, and the required parameters in this study.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Resolution/Scale</th>
<th>Source</th>
<th>Required Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEM</td>
<td>90 m</td>
<td>SRTM [56]</td>
<td>Topographical data</td>
</tr>
<tr>
<td>Soil map</td>
<td>1 km</td>
<td>FAO [41]</td>
<td>Soil classes and physical properties</td>
</tr>
<tr>
<td>Land use maps</td>
<td>60 m (1970s)</td>
<td>Landsat pre-Collection Level-1 [47], Landsat TM, ETM+, OLI Surface Reflectance Level-2, SRTM [56]</td>
<td>Land cover and use classes</td>
</tr>
<tr>
<td>Precipitation</td>
<td>Daily</td>
<td>RBWB, University of Dar es Salaam (UDSM), Tanzania Meteorological Agency (TMA)</td>
<td>Rainfall</td>
</tr>
<tr>
<td>Climate</td>
<td>Daily /0.44°</td>
<td>CORDEX Africa [45]</td>
<td>Temperature, humidity, solar radiation, wind speed</td>
</tr>
</tbody>
</table>

Table 2. Overview of the applied Regional Climate Models (RCMs) and their driving Global Climate Models (GCMs).

<table>
<thead>
<tr>
<th>GCM</th>
<th>RCM</th>
<th>Institution</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CanESM2</td>
<td>CanCM4_r2</td>
<td>Canadian Centre for Climate Modelling and Analysis (CCMa)</td>
<td><a href="http://climate-modelling.canada.ca/">http://climate-modelling.canada.ca/</a></td>
</tr>
<tr>
<td>CanESM2</td>
<td>RCA4_v1</td>
<td>Rosby Centre, Swedish Meteorological and Hydrological Institute (SMHI)</td>
<td><a href="https://esg-dn1.nsc.liu.se/">https://esg-dn1.nsc.liu.se/</a></td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>CCLM4-8-17_v1</td>
<td>Climate Limited-area Modelling Community (CLMcom)</td>
<td><a href="https://esg-dn1.nsc.liu.se/">https://esg-dn1.nsc.liu.se/</a></td>
</tr>
<tr>
<td>EC-EARTH</td>
<td>CCLM4-8-17_v1</td>
<td>Climate Limited-area Modelling Community (CLMcom)</td>
<td><a href="https://esg-dn1.nsc.liu.se/">https://esg-dn1.nsc.liu.se/</a></td>
</tr>
<tr>
<td>EC-EARTH</td>
<td>RCA4_v1</td>
<td>Rosby Centre, Swedish Meteorological and Hydrological Institute (SMHI)</td>
<td><a href="https://esg-dn1.nsc.liu.se/">https://esg-dn1.nsc.liu.se/</a></td>
</tr>
<tr>
<td>MIROC5</td>
<td>RCA4_v1</td>
<td>Rosby Centre, Swedish Meteorological and Hydrological Institute (SMHI)</td>
<td><a href="https://esg-dn1.nsc.liu.se/">https://esg-dn1.nsc.liu.se/</a></td>
</tr>
</tbody>
</table>
Figure 3. Land use and land cover classifications for four time steps ranging from (a) 1970s, (b) 1994 and (c) 2004 to (d) 2014 (modified after Leemhuis et al. [30]).

2.3. Model Description

In this study, the SWAT model [11] was applied to simulate the water balance for the chosen time period and under changing LULC conditions. SWAT is a semi-distributed and physically-based catchment model for continuous simulations of discharge, sediments, nutrients, and pesticides on a daily basis. The model divides the catchment into subcatchments, which are generated from drainage patterns derived from the DEM, and by setting a threshold that defines the minimum drainage area to form a stream. These subcatchments are further discretized into hydrologic response units (HRU) with unique combinations of LULC classes, soil types, and slope classes. LULC classes, soil types, and slope classes covering less than 10% of the area within the single subcatchments were neglected within the HRU generation. The model is divided into land phase and channel processes. Most of the processes within the land phase are simulated on the HRU level and summed up for each subcatchment, to calculate the overall water balance with the integration of climate station data and the channel processes [63]. The most important processes simulated by the model are surface runoff, infiltration, lateral flow, baseflow, evapotranspiration, and groundwater recharge. Precipitation can either be intercepted by plants or hit the ground, where it may flow as surface runoff to the reach [64],
infiltrate into the soil, or evaporate from the ground [65]. If the water infiltrates into the soil, it is stored as soil moisture or percolates with a storage routing technique, which is based on the saturated hydraulic conductivity and the field capacity of the soil profile. Lateral flows are simulated with a kinematic storage model [66]. Once water percolates below the unsaturated zone, it reaches the shallow aquifer, which is treated as an unconfined aquifer. Once a certain threshold defined by the modeler is exceeded, baseflow contributes to the reach. Water may also move into a deeper confined aquifer, where the water is assumed to contribute to the discharge outside of the catchment and is treated as lost for the processes inside the catchment. Furthermore, the water can move from the shallow aquifer into the unsaturated zone, where it can be lost by evapotranspiration. This capillary rise and evapotranspiration are controlled by the water demand of the LULC and several parameters specified by the modeler. A detailed description of the model is given by Arnold et al. [11] and Neitsch et al. [63], and further information on the model parameters can be found in Arnold et al. [67].

2.4. Model Setup and Evaluation

In total the catchment was divided into 95 subcatchments and 1087 HRUs. Eight different soil types, seven LULC classes, and five slope classes were considered (Table 3). Due to the complex topography and the high relief energy, five elevation bands [63] were integrated to account for orographic precipitation patterns, as well as altitudinal temperature changes. These elevation bands divide each subcatchment into five elevation zones. Within these zones, precipitation and temperature are modified according to the altitudinal difference among the elevation of the nearest rainfall, respective climate station, and the average elevation of the elevation band. The exact modification is calculated with a certain factor called PLAPS or TLAPS, respectively (Table 4). Evapotranspiration was calculated after Penman–Monteith [65], using historical runs of an ensemble mean of CORDEX Africa [45] data from six different models (Table 2), with a spatial resolution of 0.44° and 21 stations (Figure 1). Surface runoff and infiltration were calculated using the Soil Conservation Service (SCS) curve number method [64]. After setting up the model within ArcSWAT 2012 (revision664), the model was calibrated and validated using the SUFI-2 algorithm in SWAT-CUP (version 5.1.6.2) [68], basically following the guidelines of Arnold et al. [69] and Abbaspour et al. [70]. As the SUFI-2 program within the SWAT-CUP software was utilized for parameter optimization, the Latin Hypercube sampling iteratively discarded the worst simulations by rejecting the 2.5th and 97.5th percentile of the cumulative distribution. Therefore, the best 95% of simulations generated a parameter range (95% prediction uncertainty, 95PPU) instead of a single final parametrization. This uncertainty band represented by the 95PPU was used to account for the modeling uncertainty [68], and is quantified as the P-factor, which measures the ability of the model to bracket the observed hydrograph with the 95PPU. Finally, the P-factor is simply the fraction enveloped by the 95PPU. Hence, the P-factor can be between 0 and 1, where 1 means a 100% bracketing of the measured data. The width of the 95PPU is calculated by the R-factor (Equation (1)). The R-factor divides the average distance between the lower and upper percentile with the standard deviation of the measured data [68]. The R-factor ranges from 0 to infinity, and should be below 1, implying a small uncertainty band [68]. The final parameter ranges are illustrated in Table 4, and a detailed description of the single parameters is given in Arnold et al. [71]. The Kling–Gupta efficiency (KGE) (Equation (2)) [72] was chosen as objective function. Ancillary criteria to assess the quality of the model were the Nash–Sutcliffe Efficiency (NSE) (Equation (3)) [73], coefficient of determination ($R^2$) (Equation (4)), percent bias (PBIAS) (Equation (5)) [74], standard deviation of measured data (RSR) (Equation (6)), and the abovementioned P-factor and R-factor. Additionally, precipitation distribution was assessed with remote sensing products [38,75], a baseflow filter technique [76] was utilized to estimate the share of surface runoff and baseflow, and literature research was performed [24–27,36,40,77,78] to evaluate the gathered water balance component values for plausibility. After calibration and validation, different land use maps (Figure 3) were utilized to simulate the impact of LULCC. In order to attribute all alterations to the LULCC, nothing was modified except for the land use maps.
Table 3. Final spatial coverage with the different soil types, LULC classes, and slope classes of the study area after defining the HRUs.

<table>
<thead>
<tr>
<th>Soil Type</th>
<th>Proportional Area (%)</th>
<th>LULC Class (1970s Land Use)</th>
<th>Proportional Area (%)</th>
<th>Slope Classes</th>
<th>Proportional Area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplic Acrisol</td>
<td>30.82%</td>
<td>Savanna</td>
<td>45.30%</td>
<td>0–2%</td>
<td>22.45%</td>
</tr>
<tr>
<td>Eutric Fluvisol</td>
<td>14.77%</td>
<td>Range-Grasses</td>
<td>23.49%</td>
<td>2%–5%</td>
<td>12.09%</td>
</tr>
<tr>
<td>Humic Nitisol</td>
<td>13.14%</td>
<td>Forest-Mixed</td>
<td>22.05%</td>
<td>5%–8%</td>
<td>11.66%</td>
</tr>
<tr>
<td>Ferric Lixisol</td>
<td>13.05%</td>
<td>Forest-Evergreen</td>
<td>6.81%</td>
<td>8%–12%</td>
<td>13.96%</td>
</tr>
<tr>
<td>Ferric Acrisol</td>
<td>9.08%</td>
<td>Agricultural Land</td>
<td>1.20%</td>
<td>&gt;12%</td>
<td>39.85%</td>
</tr>
<tr>
<td>Humic Acrisol</td>
<td>5.73%</td>
<td>Water</td>
<td>0.95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albic Arenosol</td>
<td>2.75%</td>
<td>Barren</td>
<td>0.20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eutric Leptosol</td>
<td>0.59%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Ranking of the calibrated parameters, according to their sensitivity and significance. A “v” in Method implies a replacement of the initial parameter value with the given value in the final range, whereas an “r” indicates a relative change to the initial parameter value.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Parameter</th>
<th>Description</th>
<th>Final Range</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GWQMN.gw</td>
<td>Threshold depth of water in the shallow aquifer for return flow to occur (mm).</td>
<td>1400–2200</td>
<td>v</td>
</tr>
<tr>
<td>2</td>
<td>ALPHA_BF.gw</td>
<td>Base flow alpha factor (days).</td>
<td>0.15–0.26</td>
<td>v</td>
</tr>
<tr>
<td>3</td>
<td>GW_REVAP.gw</td>
<td>Groundwater “revap” coefficient.</td>
<td>0.15–0.2</td>
<td>v</td>
</tr>
<tr>
<td>4</td>
<td>SURLAG.bsn</td>
<td>Surface runoff lag coefficient.</td>
<td>2.8–5.3</td>
<td>v</td>
</tr>
<tr>
<td>5</td>
<td>GW_DELAY.gw</td>
<td>Groundwater delay time (days).</td>
<td>4–30</td>
<td>v</td>
</tr>
<tr>
<td>6</td>
<td>SOL_K().sol</td>
<td>Saturated hydraulic conductivity (mm/h).</td>
<td>0.4–0.7</td>
<td>r</td>
</tr>
<tr>
<td>7</td>
<td>RCHRG_DP.gw</td>
<td>Deep aquifer percolation fraction.</td>
<td>0.31–0.37</td>
<td>v</td>
</tr>
<tr>
<td>8</td>
<td>SOL_Z().sol</td>
<td>Depth from soil surface to bottom of layer (mm).</td>
<td>0.35–0.5</td>
<td>r</td>
</tr>
<tr>
<td>9</td>
<td>SOL_AWC().sol</td>
<td>Available water capacity of the soil layer (mm H₂O/mm soil).</td>
<td>–0.1–0.13</td>
<td>r</td>
</tr>
<tr>
<td>10</td>
<td>R_OV_N.hru</td>
<td>Manning’s “n” value for overland flow.</td>
<td>0.1–0.2</td>
<td>r</td>
</tr>
<tr>
<td>11</td>
<td>R_CN2.mgt</td>
<td>SCS runoff curve number for moisture condition II.</td>
<td>–0.5–(−0.35)</td>
<td>r</td>
</tr>
<tr>
<td>12</td>
<td>CH_K1.sub</td>
<td>Effective hydraulic conductivity in the tributary channel (mm/h).</td>
<td>65–80</td>
<td>v</td>
</tr>
<tr>
<td>13</td>
<td>ESCO.hru</td>
<td>Soil evaporation compensation factor.</td>
<td>0–0.1</td>
<td>v</td>
</tr>
<tr>
<td>14</td>
<td>CH_K2.rte</td>
<td>Effective hydraulic conductivity in the main channel (mm/h)</td>
<td>100–130</td>
<td>v</td>
</tr>
<tr>
<td>15</td>
<td>REVAPMNN.gw</td>
<td>Threshold depth of water in the shallow aquifer for “revap” to occur (mm).</td>
<td>13–30</td>
<td>v</td>
</tr>
<tr>
<td>16</td>
<td>EPCO.hru</td>
<td>Plant uptake compensation factor.</td>
<td>0.9–1</td>
<td>v</td>
</tr>
<tr>
<td>17</td>
<td>PLAPS</td>
<td>Precipitation lapse rate (mm H₂O/km).</td>
<td>130</td>
<td>v</td>
</tr>
<tr>
<td>18</td>
<td>TLAPS</td>
<td>Temperature lapse rate (°C/km).</td>
<td>–6</td>
<td>v</td>
</tr>
</tbody>
</table>

\[
R - factor = \frac{1}{n\sigma_O} \sum_{n=1}^{n} (S_U - S_L)
\]  

\[
KGE = 1 - \sqrt{(r - 1)^2 + (a - 1)^2 + (\beta - 1)^2}
\]  

\[
NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}
\]  

\[
R^2 = \frac{\left[\sum_{i=1}^{n}(O_i - \bar{O})(S_i - \bar{S}) \right]^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(S_i - \bar{S})^2}
\]  

\[
PBIAS = 100 \times \frac{\sum_{i=1}^{n}(O_i - S_i)}{\sum_{i=1}^{n} O_i}
\]
\[ RSR = \sqrt{\frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(\bar{O} - \bar{S})^2}} \]

where \( n \) is the number of observations, \( c_\sigma \) is the standard deviation of the observed discharge, with \( S_{U} \) and \( S_{L} \) representing the upper 97.5th and lower 2.5th percentiles of the simulated 95PPU, respectively; \( r \) is the linear regression coefficient between observed and simulated data; \( \alpha \) is the ratio of the standard deviation of simulated and observed data; \( \beta \) is the ratio of the means of simulated and observed data; \( O_i \) and \( S_i \) are the observed and simulated discharge values, respectively; and \( \bar{O} \) and \( \bar{S} \) are the mean of observed and simulated discharge values.

3. Results

3.1. Model Calibration and Validation

Figure 4 illustrates the results of the calibration and validation for the modeled discharge, using the SUFI-2 calibration technique. The hydrograph generally indicates a good fit of the daily discharge dynamics by the SWAT model. This is also emphasized by the statistical quality of the model shown in Table 5. According to Moriasi et al. [79], PBIAS, NSE, and \( R^2 \) perform very well for both the calibration and the validation period. Nevertheless, in some years (e.g., 1959, 1961) the model overestimates discharge, whereas in contrast, slight underestimations can be observed at the transitions from the dry to the rainy seasons (e.g., 1963/1964, 1964/1965, 1965/1966, 1966/1967, 1969/1970) (Figure 4). However, the general model performance shows a good agreement between simulated and observed discharge, which is also highlighted by the flow duration curve (FDC) (Figure 5). The FDC nearly indicates a perfect fit, with slight underestimations of the low flows and the upper 2%–3% of the exceedance probabilities by the model. These extreme flows account in total for about 11%–15% of the annual water yield. The uncertainty of the simulations represented by the 95PPU band are quite low for the low flows, whereas the uncertainty is highest for flows between 1000–2300 m^3/s. This can be attributed to some overestimated peaks (1959, 1961) and also the model’s difficulties in simulating a small peak just before the main peak (e.g., 1963, 1964). The water balance values (Table 6) are consistent with other publications in this area with regard to the groundwater recharge [36,80], precipitation [24,27,38], evapotranspiration [78], and potential evapotranspiration [81]. The ratio of surface runoff and baseflow coincides with the baseflow filter technique of Arnold et al. [76]. Precipitation data is the key driver of rainfall runoff models, but is also a source of uncertainty, especially in data-scarce regions [82]. Due to the low number of precipitation stations and their distribution within the catchment in either low altitudes in the eastern part or high altitudes in the western part of the catchment, elevation bands [63] were implemented. The effect on the average precipitation among the subcatchments is shown in Figure 6. The figure clearly illustrates the variation of precipitation within the single subcatchments. This is particularly true in the eastern part, where the mountainous region of the Udzungwa Mountains in the north and the Mahenge Highlands in the south receive more than 50 mm of additional precipitation as measured in the valley. In the western part of the catchment, a reverse effect is visible, with decreased precipitation in the low altitude subcatchments due to the high altitude of the precipitation stations. All these changes are attributed to the implementation of the elevation bands, as all parameters of the final parameter solution are unchanged except for the elevation bands. Due to the numerous solutions within the uncertainty band, this is only one representative example of the importance of orographic precipitation patterns in the Kilombero Catchment.
Table 5. Summary of the quantitative model performance analysis for the calibration and validation period. P-factor is the percentage of measured data covered by the 95PPU uncertainty band, R-factor is the relative width of the 95PPU uncertainty band, $R^2$ is the coefficient of determination, NSE is the Nash–Sutcliffe efficiency, PBIAS is the percent bias, KGE is the Kling–Gupta efficiency, and RSR the standard deviation of measured data.

<table>
<thead>
<tr>
<th>Simulation Period (Daily)</th>
<th>P-Factor</th>
<th>R-Factor</th>
<th>$R^2$</th>
<th>NSE</th>
<th>PBIAS</th>
<th>KGE</th>
<th>RSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration (1958–1965)</td>
<td>0.62</td>
<td>0.45</td>
<td>0.86</td>
<td>0.85</td>
<td>0.3</td>
<td>0.93</td>
<td>0.38</td>
</tr>
<tr>
<td>Validation (1966–1970)</td>
<td>0.67</td>
<td>0.56</td>
<td>0.80</td>
<td>0.80</td>
<td>2.5</td>
<td>0.89</td>
<td>0.45</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Water Balance Components</th>
<th>in (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>1344</td>
</tr>
<tr>
<td>Actual evapotranspiration</td>
<td>788</td>
</tr>
<tr>
<td>Potential evapotranspiration</td>
<td>1380</td>
</tr>
<tr>
<td>Surface runoff</td>
<td>43</td>
</tr>
<tr>
<td>Lateral flow</td>
<td>58</td>
</tr>
<tr>
<td>Base flow</td>
<td>209</td>
</tr>
<tr>
<td>Recharge to the deep aquifer</td>
<td>242</td>
</tr>
</tbody>
</table>

Figure 4. Hydrograph showing the observed and the simulated discharge for the calibration (1958–1965) and the validation period (1966–1970), separated by the vertical dashed line. Statistical measures are shown in Table 5.
Figure 5. Flow duration curve for the simulated and observed discharge for the simulation period (1958–1970). Additionally, the 95PPU is illustrated, which represents the modeling uncertainty by showing the cumulative distribution of flow between the 2.5th and 97.5th percentiles of all the simulation solutions. Statistical measures for the ranked simulated and observed values are 0.99 for $R^2$ and NSE, respectively.

Figure 6. Spatial variations in average precipitation (mm) between 1958–1970, due to the implementation of elevation bands. Positive values correspond to increased precipitation due to the implementation of elevation bands, and vice versa.

3.2. Spatio-Temporal Analysis

Precipitation is the main driver for hydrological processes [83], hence an overview of the temporal precipitation pattern (Figure 7) is crucial for the interpretation of the hydrological processes occurring in the catchment. Figure 7 underlines the distinction of wet and dry seasons by showing the monthly precipitation patterns for the entire catchment. The rainy season starts in November/December and lasts until April—and in some years up to May (1961, 1967, and 1968). Alterations with bimodal and unimodal patterns are also visible, whereas some years show pronounced monthly precipitation peaks in December or January, representing the irregular occurrence of the small rainy season (1961, 1967, and 1970) [38]. In March 1963, which was an exceptionally wet period over all East Africa [77], precipitation exceeded 400 mm.
The boxplots in Figure 8 illustrate the high (a) interannual and (b) intraannual variability of discharge. The clear distinction between the wet and dry seasons is obvious from the boxplots in Figure 8b. The seasonal distinction of the overall discharge is deconstructed into the single water balance components on a monthly timescale in Figure 9. Lateral and surface runoff only occur in the rainy season, with peaks in March and April, whereas the more pronounced baseflow peaks in April and May result in the highest water yield in April. Evapotranspiration and potential evapotranspiration are almost identical from February to May, but differ by more than 150 mm in September and October, indicating a water deficit in the dry season. The potential evapotranspiration more or less follows an antithetical pattern compared to the precipitation pattern (Figure 9).
Figure 8. Boxplots for discharge on (a) daily and (b) monthly resolution, emphasizing the variation of discharge on different timescales and the distinction of wet and dry season.

Figure 9. Monthly averages of the water balance components for the Kilombero Catchment within the simulation period (1958–1970).
Figure 10 illustrates the influence of single water balance components on the subcatchment scale. The spatial pattern of the water yield fluxes (Figure 10a–d) mirrors the general picture of the precipitation, by showing a high contribution from the mountainous subcatchments compared to the contribution of water from the valley. The patterns of surface (Figure 10a), lateral (Figure 10b), and groundwater (Figure 10c) contribution also show contradictory patterns in several subcatchments. Surface runoff and lateral flow show some discharge hotspots, which mainly occur in steep areas with clayey (surface runoff) or sandy (lateral flow) topsoils. Figure 10d shows the overall water yield, underlining the contribution of water from the mountains to the valley, especially from the northern Udzungwa Mountains and the Mahenge Highlands. Vice versa, the evapotranspiration and potential evapotranspiration patterns exhibit high values in the valley compared to the mountainous parts of the catchment (Figure 10e,f).

Figure 10. Spatial mapping of mean annual values of (a) surface runoff contribution, (b) lateral flow contribution, (c) groundwater contribution, (d) the overall water yield, (e) actual evapotranspiration, and (f) the potential evapotranspiration for the subcatchments.

3.3. Land Use and Land Cover Changes and Their Impact on Water Resources

Figure 11 shows the percentage shifts among the LULC classes from the 1970s up to 2014, as opposed to the classes’ spatial representation shown in Figure 3. In this case study, most parts of the wetland are classified as grassland and cropland, as it is a seasonally flooded grassland prone to conversions into cropland (Figure 3a–d). Noteworthy is the high share of savanna and grassland, as well as the shift among savanna, grassland, and agriculture, besides forest-mixed and forest-evergreen, which is mainly attributed to problems in the classification of the Landsat images. These misclassifications are mainly caused by spectral class similarity and lack of suitable data, especially in early dates. Nevertheless, the share of agricultural land increased significantly from 2004 to 2014, which is also reported by other studies [30,33]. The fringe of the wetland and the western part of the catchment (Figure 3) are most strongly affected by an increase in agricultural land.
Figure 11. Percentage share of the land use/land cover (LULC) classes within the Kilombero Catchment from the 1970s up to 2014.

Figure 12 shows the variations of surface runoff (Figure 12a), groundwater contribution (Figure 12b), evapotranspiration (Figure 12c), and the overall water yield (Figure 12d) on subcatchment scale between the LULC setup from the 1970s to 2014. Evapotranspiration and groundwater contribution show decreasing trends in the floodplain, where grassland is turned into cropland and the share of water on the total land cover is reduced (Figures 3 and 11). One subcatchment in the western part, where grassland was converted into either agricultural land or barren, indicates a strong decrease in groundwater contribution (Figures 3 and 12b). This decrease in groundwater contribution coincides with higher evapotranspiration (agricultural land) and higher surface runoff (barren). However, many subcatchments in the mountainous parts show increasing groundwater contributions and water losses due to evapotranspiration. These higher values are mainly occurring in subcatchments with a rising share of evergreen forests or savanna in 2014. Between 1970 and 2014, in areas where savanna was converted into grassland or agricultural land, the development of surface runoff showed an inverted development, with increasing surface runoff amounts. This happened mainly in the eastern parts of the catchment and on the fringe of the floodplain (Figure 12a). The general picture indicates a decrease of water yield (Figure 12d) in the Kilombero Valley related to the lower evapotranspiration values, whereas some subcatchments in the Udzungwa Mountains and the Mahenge Highlands show increasing water contribution, which corresponds to the spatial pattern of the increased groundwater contribution (Figure 12b).
Figure 12. Average shifts in water balance components (in mm) for the simulation period, comparing changes from the 1970s land use map with the 2014 land use map. (a) Deviations in surface runoff, (b) groundwater contribution, (c) evapotranspiration, and (d) the overall water yield are displayed.

Figure 13 shows the average hydrological impacts within the simulation period for the Kilombero Catchment, considering another temporal scale that shows the monthly changes for the single water balance components. Except for increasing groundwater flow and changes in evapotranspiration (Figure 13b), the monthly changes within the entire catchment are rather small. The shift from 2004 to 2014 (Figure 13d) seems to be negligible. However, Figure 14 shows more pronounced the effects of the LULCC from 2004 to 2014 on the subcatchment scale. Surface runoff contribution is increasing in almost the entire valley and in the eastern Udzungwa Mountains by up to 10 mm (Figure 14a), which is 23% higher surface runoff compared to the average catchment surface runoff (Table 6). This is due to accelerated conversion into agricultural land. In contrast, the groundwater contribution is decreasing by up to 20 mm within this area, reinforcing changes in the system’s hydrology (Figure 14b). The overall water yield patterns (Figure 14d) are more complex, with decreasing water fluxes in subcatchments prone to anthropogenic activities in the fringe of the wetland, due to the lower groundwater contribution. Hence, water yield in the upper western part of the catchment is increasing, because of the increasing surface runoff and the conversions into barren and cropland. Evapotranspiration (Figure 14c) is slightly increasing within the wetland and in most of the mountainous subcatchments, resulting in a loss of water for the catchment. Only the fringe of the wetland shows coherent lower evapotranspiration values. The changes from 2004 to 2014 were exemplarily chosen to be shown here, as their pattern of change with the distinct conversion into cropland (Figure 11) is the most probable future land use pattern for the Kilombero Catchment.
Figure 13. Shifts in water balance components (in mm) for the entire catchment on a monthly time scale running the model with four different land use maps. (a) Shows the water balance of the 1970s land use map run, (b) displays a comparison of the 1970s map with the land use of 1994, (c) the changes from 1994 to 2004, and (d) illustrates the shifts from 2004 to 2014. All inputs except for land use maps are not modified.

Figure 14. Average shifts in water balance components (in mm) for the simulation period, while comparing changes from the 2004 land use map with the 2014 land use map. (a) Changes in surface runoff, (b) groundwater contribution, (c) evapotranspiration, and (d) the overall water yield.
4. Discussion

4.1. Model Evaluation and Spatio-Temporal Analysis

Despite the low number of precipitation stations, the spatio-temporal precipitation pattern of
the catchment is represented quite well with the implementation of elevation bands, which was
crucial due to the high altitude of the precipitation stations in the western part and the low altitude
of the stations in the eastern part of the catchment (Figure 6). A comparison of global precipitation
datasets (GPDS) has already been published by Koutsouris et al. [38], visualising the average spatial
precipitation distribution for the Kilombero Catchment among frequently-utilized GPDS. As a result
of their study, Koutsouris et al. [38] showed large differences in the spatial precipitation patterns
of eight different products in some areas within the catchment. What all GPDS have in common is
the relatively high rainfall in the Udzungwa Mountains and Mahenge Highlands, which was also
the case for the precipitation stations utilized in this study, due to the orographic correction factor
implemented with the elevation bands (Figure 6). It should not be ignored that most of the GPDS
estimated relatively high precipitation amounts for the southwestern part of the catchment, which
contradicts with station data available within this study. However, it should be noted that the study
of Koutsouris et al. [38] applied satellite precipitation estimates that were not operational during our
period of investigation, whereas our utilized precipitation stations stopped being functional before the
onset of the satellite products. This temporal mismatch aggravated further comparisons among the
GPDS and station data, and we concentrated on the general spatial patterns of precipitation. Moreover,
the GPDS pattern for the southwestern region can just be altered by ignoring all three available
precipitation stations in this area. This is also not a suitable option, keeping in mind the generally high
uncertainty with regard to precipitation patterns in that region, according to the GPDS, which has
already been proven by Koutsouris et al. [38]. For example, the difference between the station data
and the patterns of the GPDS in the southwestern catchment area is still smaller than the differences
between certain GPDS, like the Global Precipitation and Climatology Center v6 data set (GPCC) [84]
and the Modern Era Retrospective-Analysis for Research and Applications (MERRA) [85]. These both
show opposing results when comparing precipitation patterns from the Mahenge Highlands and the
Udzungwa Mountains [38], and therefore much larger differences than the station data and the general
picture of the GPDS in the southwestern Kilombero Catchment. Thus, in mountainous tropical regions,
with persistent cloud coverage, an understanding of the strengths and limitations of remote sensing
products are a prerequisite for an adequate application [6]. In spite of these uncertainties, large-scale
precipitation patterns are captured quite well with remote sensing products [21,86].

The hydrograph (Figure 4), the flow duration curve (Figure 5), and the statistical model
performance (Table 5) all indicate a good to very good performance of the SWAT model for the
simulation period [79]. However, some peaks are not captured well, and for some years the discharge
is underestimated (1959, 1961 in Figure 4), which leads to the slightly lower NSE (Table 5) compared to
the other evaluation criteria, because of the high sensitivity of the NSE to peaks [87]. For the model
setup, this study followed the procedure and calibration techniques given by Arnold et al. [69] and
Abbaspour et al. [70], by keeping the model parametrization within the requirements of parsimony
and robustness [88]. Although the number of parameters is quite large compared to other model
applications in the tropics [89], it is still in the range of similar applications of the SWAT model
under tropical conditions [15,90–92]. Five out of the seven most sensitive parameters are related
to groundwater (Table 4), which underlines the importance of baseflow for the catchments water
yield. These parameters are typical parameters used in SWAT to calibrate baseflow, which was
also demonstrated in a meta study by Arnold et al. [69]. In conclusion, these parameters mainly
control the occurrence (GWQMN), recession (ALPHA_BF, GW_DELAY), and the vertical movement
of groundwater (GW_REVAP, RCHRG_DP), and were calibrated within the default ranges given
by SWAT-CUP. The relevance of baseflow in the Kilombero Catchment was already highlighted by
Burghof et al. [26]. Figure 9 illustrates that baseflow contributes nearly 100% of the water yield from
June to November on the catchment scale. Furthermore, Gabiri et al. [36] and Burghof et al. [93] showed that the depth to the groundwater table is closer to the surface at the fringe of the floodplain compared to the riparian zone, due to the high influence of baseflow contribution from the mountains, especially in the dry season. This shallow groundwater affects plant growth patterns and agricultural activities within the valley, and transects from the river to the fringe with year-round water availability for deeply rooted plants (1–2 m) in the fringe of the floodplain [36]. With regard to water management and agricultural utilization of the floodplain [37], the findings of this research combined with the aforementioned information on groundwater contribution [26,36] might raise awareness for the importance of the upland catchment, which is closely linked to the wetland system. This linkage is represented by the already highlighted influence of year-round groundwater contribution from the higher elevations into the valley bottom. Model results show that groundwater contribution is virtually the only water source from June to November (Figure 9), and this groundwater is generated in the upper catchment (Figure 10d), whereas the wetland itself is prone to high evapotranspiration (Figure 10e) and contributes much less water to the stream (Figure 10d).

4.2. Impact of Land Use and Land Cover Change

Despite different technologies, there is high congruence between the Landsat 5, 7, and 8 sensors, which have the same spatial resolution of 30 m. Their band definitions differ only slightly for most bands, and the effect has been found to be negligible [94]. A stronger technical and methodological difference exists between the 1970s time step, using the Landsat pre-collection Level 1 at 60 m spatial resolution, and using a conventional mosaicking method due to the lack of a sufficient number of images. The post-classification comparison (PCC) method was used for detecting change, as methodological and technological inter-classification differences are less important. More crucial is the respective classification accuracy, as with PCC errors are propagated [95]. For the Kilombero Catchment, classes more prone to error were the natural classes: savanna, range grasses, wetland, and forest-mixed. Some confusion also exists among savanna, range grasses, and agriculture; however, our PCC results are mostly logically consistent and conform to historical maps. The conversion of natural classes to agriculture results in rather strong spectral changes, whereas the modification of forests by single tree extraction cannot be adequately resolved, neither with PCC nor with methods based on spectral bands.

Due to the aforementioned circumstances, the conversion of forested areas and savanna in the upper catchment might significantly influence water quantity and the year-round water contribution to the stream. Deforestation activities in the entire catchment are already occurring [30,32], and may lead to a shift from slow groundwater contribution to fast surface water contribution from the uplands [96]. The increased share of cropland, which results in a reduced retention capacity, will influence the flow regime, with declining low flows and aggravated flooding. This is especially important, as vulnerability to floods and droughts is already highlighted as a challenge for the floodplain area in the Integrated Water Resources Management and Development Plan (IWRMDP) [97]. The conversion of forests and savanna into crop- and grassland and its subsequent hydrological impacts will aggravate this vulnerability. These relationships are visualized in Figures 13 and 14. While interpreting Figure 13, one has to keep in mind that complex large river basins like the Kilombero could conceal small-scale effects [98], as already shown by Wagner et al. [99]. Furthermore, this figure illustrates monthly averages of a 13-year simulation period, implying additional concealing effects and a too-broad time scale to account for daily events. Following the approach of Wagner et al. [99], Figure 14 shows the impacts on subcatchment scale, with increasing surface runoff due to anthropogenic activities. The decreasing evapotranspiration at the fringe of the floodplain (Figure 14c) can be attributed to the lower evapotranspiration of the agricultural land. For the envisioned large-scale rice schemes of the Southern Agricultural Growth Corridor of Tanzania (SAGCOT) plans, the floodplain could be modified into rice instead of agricultural land use, which will influence evapotranspiration significantly. Figure 14 generally shows a complex picture with regard to changes in water balance.
components (see Section 3.3). This complex picture fosters the concealing effects of the large catchment. Furthermore, the different results from Figures 13 and 14, as well as the temporal changes in water balance components (Figure 9) underline the scale dependency of the hydrological processes in both space and time within the Kilombero Catchment, and therefore the need to consider various spatio-temporal scales for water management plans.

Apart from these scale effects, LULCC within the valley itself in relation to the implementation of a growth corridor are far from being negligible concerning water quantity. The especially high contribution of groundwater fluxes from the upper catchment throughout the whole year is important for the wetland, its vegetation, and also for the agricultural activities and the attached food security. The implicated conversion from grassland to cropland in the growth corridor [30] additionally affects water quality negatively, which is another important aspect when investigating wetlands. Regarding the transport of sediments, nutrients, pesticides, and bacteria, data availability is insufficient. Nevertheless, this topic could be interesting for future investigations, especially with regard to the planned intensification of agricultural activities [37]. These increased agricultural activities potentially result in economic benefits, consequently followed by increasing population through demographic growths and migration. This increasing population might lead to further encroachment of the uplands, and therefore increased pressure on savanna as well as upland forests, which will foster the aforementioned changes with regard to water resources. Considering these circumstances, the long-term effects might therefore imply increasing surface runoff contributions due to upland deforestation, and consequently lower retention potentials for flood mitigation in the rainy season, as well as decreased low flow supply in the dry season. Another aspect that needs more investigation from social science perspective, with regard to the planned large-scale utilization of the valley, are rising land use conflicts among farmers and pastoralists, which are often caused by a lack of sufficient pasture or water supplies [100].

According to the IWRMDP, significant groundwater resources exist within the northeastern part of the Ramsar site (Kibasila Wetland). These groundwater resources are seen as a potential source for irrigation, although the overall potential of this aquifer is not yet explored and the implications on surface-groundwater interactions are uncertain [97]. However, for some areas, a moderately low hydraulic conductivity towards the riparian zone has already been investigated [36,101]. For a sustainable use of water resources in the Kilombero Catchment, future research should focus on the ongoing responses to anthropogenic impacts like land cover conversion and the impact of climate change, as well as management impacts, such as the construction of dams or the intensified utilization of nutrients and pesticides.

5. Conclusions

A combination of local discharge and precipitation data, combined with multi-temporal Landsat images and freely available geo datasets, allowed a detailed and distributed analysis of the hydrological system of a topographical complex East African catchment. This is the first study with distributed information on the water balance in Kilombero Catchment, which is strongly affected by LULCC and will be further affected by climate change and more pronounced LULCC in the near future. As it also comprises a Ramsar site, many interests collide within the catchment. They need to be harmonized by sustainable water management, and therefore, well-informed decisions are needed. This study showed the scale dependency of water resources in the Kilombero Catchment and the need for distributed modeling. It was demonstrated that the wetland is severely dependent on mountainous water resources and year-round groundwater contribution. Therefore, we emphasize the necessity of protecting upland forests as one important factor to ensure a perennial water supply for the valley bottom, its embedded wetland, and the inherent ecosystem services provided by both the wetland and the upland forests. So far, LULCC occur predominantly within the Kilombero Valley, and has had rather local effects on the water balance components over the past years. At the same time, the mountainous areas that are the most important source of groundwater experienced much less
LULCC. For future management of the Kilombero Catchment, it will be important to prevent these upland areas from extensive LULCC, in order to sustain water availability in the wetland. Apart from this case study, this article might serve as an example of how to utilize the available historic precipitation [102] and discharge data [103], and how to combine them with historic climate model runs, global soil data, and LULC data gathered from earth observation images for meso- to large-scale applications, especially in data-sparse regions like Sub-Saharan Africa.

**Author Contributions:** K.N., B.D. and C.L. conceived and designed the experiments; K.N. performed the experiments; K.N., B.D. and C.L. analyzed the data; S.S. and F.T. produced the land use maps; R.v.d.L. and L.S.S. processed the climate data. K.N. wrote the article. All authors made revisions and improvements to the final version.

**Funding:** This study was supported through funding from the German Federal Ministry of Education and Research (FKZ: 031A250A-H); German Federal Ministry for Economic Cooperation and Development under the “GlobE: Wetlands in East Africa” project.

**Acknowledgments:** The authors would like to thank the Rufiji Basin Water Board for data sharing and assistance in the field and Salome Misana for her administrative guidance and logistic assistance.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

**References**

17. Wagner, P.D.; Kumar, S.; Schneider, K. An assessment of land use change impacts on the water resources of the Mula and Mutha Rivers catchment upstream of Pune, India. *Hydrol. Earth Syst. Sci.* 2013, 17, 2233–2246. [CrossRef]
18. Baker, T.J.; Miller, S.N. Using the Soil and Water Assessment Tool (SWAT) to assess land use impact on water resources in an East African watershed. *J. Hydrol.* 2013, 486, 100–111. [CrossRef]
30. Leehmuis, C.; Thonfeld, F.; Näsch, K.; Steinbach, S.; Muro, J.; Strauch, A.; López, A.; Daconto, G.; Games, I.; Diekkrüger, B. Sustainability in the food-water-ecosystem nexus: The role of land use and land cover change for water resources and ecosystems in the Kilombero Wetland, Tanzania. *Sustainability* 2017, 9, 1513. [CrossRef]


43. Kato, F. Development of a major rice cultivation area in the Kilombero Valley, Tanzania. Afr. Study Monogr. 2007, 36, 3–18. [CrossRef]


46. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 2011, 137, 553–597. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).