Evaluation of Spatio-Temporal Patterns of Remotely Sensed Evapotranspiration to Infer Information about Hydrological Behaviour in a Data-Scarce Region

Frank Joseph Wambura [1,2,3,*], Ottfried Dietrich [1] and Gunnar Lischeid [1,2]

[1]   Institute of Landscape Hydrology, Leibniz Centre for Agricultural Landscape Research (ZALF) e. V.,
     Eberswalder Straße 84, D-15374 Müncheberg, Germany; odietrich@zalf.de (O.D.); lischeid@zalf.de (G.L.)
[2]   Institute of Earth and Environmental Science, University of Potsdam, Karl-Liebknecht-Str. 24–25,
     D-14476 Potsdam-Golm, Germany
[3]   Department of Urban and Regional Planning, Ardhi University, P.O. Box 35176, Dar es Salaam, Tanzania
[*]   Correspondence: fwambura@zalf.de or wamburafj@gmail.com; Tel.: +49-33432-82-350;
     Fax: +49-33432-82-301

Abstract: Information about the hydrological behaviour of a river basin prior to setting up, calibrating and validating a distributed hydrological model requires extensive datasets that are hardly available for many parts of the world due to insufficient monitoring networks. In this study, the focus was on prevailing spatio-temporal patterns of remotely sensed evapotranspiration (ET) that enabled conclusions to be drawn about the hydrological behaviour and spatial peculiarities of a river basin at rather high spatial resolution. The prevailing spatio-temporal patterns of ET were identified using a principal component analysis of a time series of 644 images of MODIS ET covering the Wami River basin (Tanzania) between the years 2000 and 2013. The time series of the loadings on the principal components were analysed for seasonality and significant long-term trends. The spatial patterns of principal component scores were tested for significant correlation with elevations and slopes, and for differences between different soil texture and land use classes. The results inferred that the temporal and spatial patterns of ET were related to those of preceding rainfalls. At the end of the dry season, high ET was maintained only in areas of shallow groundwater and in cloud forest nature reserves. A region of clear reduction of ET in the long-term was related to massive land use change. The results also confirmed that most soil texture and land use classes differed significantly. Moreover, ET was exceptionally high in natural forests and loam soil, and very low in bushland and sandy-loam soil. Clearly, this approach has shown great potential of publicly available remote sensing data in providing a sound basis for water resources management as well as for distributed hydrological models in data-scarce river basins at lower latitudes.

Keywords: evapotranspiration; hydrological behaviour; land cover change; MODIS ET; principal component analysis; shallow groundwater

1. Introduction

A sound understanding of hydrologic cycles of river basins is a crucial part in planning and managing water resources. Reliable predictions from distributed hydrological models require extensive datasets for setup, calibration and validation. Usually, time series of discharge and groundwater head are used to assess hydrological behaviour. However, these time series are often too short, corrupt and not even available in data-scarce regions. In addition, information provided by a hydrograph is integrated in space. In contrast, evapotranspiration (ET) data based on multi-temporal remote

sensing products are now increasingly available for many parts of the world, and hence can be used to provide extensive information about hydrological behaviour in river basins. ET data comprise a direct quantification of hydrological fluxes that are second to precipitation, exceeding groundwater recharge and runoff in many regions of the world. Therefore, sophisticated analysis of remotely sensed ET data might even enable us to infer spatio-temporally distributed information about boundary conditions for groundwater, climate, elevations, slopes, soil texture, land use, and long-term trends for single variables of influence which are of vital importance for management as well as hydrological modelling. This information may help reduce the number of computational units in a distributed hydrological model as well as constraining it.

In spite of the fact that ET is at the centre of hydrological processes in river basins, its conventional measurements or estimations (pan evaporation, lysimeter data, etc.) are not available at the desired spatial and temporal resolutions. Therefore, remote sensing is the only viable option to map ET of relatively large areas in a globally consistent and economically feasible manner [1,2]. Remotely sensed ET data provide only a snapshot of ET at an instance of image acquisition, but repeated images capture seasonal changes and long-term trends. The recent availability of satellite images of medium spatio-temporal resolution from the MODerate resolution Imaging Spectroradiometer (MODIS) imagery program has increased its application potential [3,4]. One of the products of this imagery program is the MODIS evapotranspiration (MODIS ET) data.

In Africa, and specifically Sub-Saharan Africa, the MODIS ET data has been widely applied [3–10]. In Tanzania particularly, the MODIS ET data has been applied in three river basins including the Wami River basin [3,9,10]. For example, the data have been applied in the Internal Drainage basin in the northwest of the Wami River basin to assess the dynamics of Lake Manyara [9]. In that study, Deus and Gloaguen [9] evaluated the suitability of the MODIS ET by comparing with measured pan ET and found coefficients of correlation of 0.60 and 0.73 for monthly and annual ET respectively. In the Pangani River basin in the north of the Wami River basin, the data were used to map ET trends [3]. In that basin, Kiptala, et al. [3] found that in vegetated land surface, the MODIS ET has a coefficient of spatial correction of 0.74 when compared with monthly ET derived from the surface energy balance procedure. In the Wami River basin, the data have also been used to assess the seasonal variation of ET [10]. These examples demonstrate that the MODIS ET data has a great potential for hydrological studies in this region.

Interpreting a long time series with high spatio-temporal variability is not a trivial task [11–13]. Exploratory data analysis using multivariate statistical techniques for dimensionality reduction, for example principal component analysis (PCA) have been used in various water-related studies [14–23]. In remote sensing studies, methods such as minimum noise fraction (MNF) and PCA have been applied for dimensionality reduction, pattern recognition, land cover change detection, land cover characterization, image transformation, image classification, and image fusion [24–32]. In this study, PCA was used for dimensionality reduction because it generates higher signal-to-noise ratio than MNF when mixed pixels are used (i.e., a mixed pixel maps several different ground objects) [23,33]. This inconvenience of different neighbouring physical areas being mapped onto the same pixel is common in MODIS data because of the whiskbroom design and the observation geometry of the MODIS instrument, and spatial uncertainty in the registration of images [34]. Moreover, Xavier et al. [35] also argue that PCA is able to reveal interannual and non-periodical signals such as long-term trend, short-term trend, and seasonal variation in the time series of remotely sensed data.

In some studies, remote sensing data have been primarily used to assess spatial patterns of variables like land cover [28,30], that could then be inputted directly into hydrological models. This study followed a different approach. An extensive dataset of multi-temporal images of ET was systematically analysed to extract as much information as possible about features that help to better understand hydrological processes in a data-scarce region. For example, enhanced ET during long dry periods might indicate regions of extensive irrigation of crops or shallow groundwater being available for plant roots, long-term shifts of ET patterns could reflect major land use changes, etc. Thus, the focus of the analysis was on spatial patterns during certain boundary conditions, or even on changes of spatial patterns. This

approach exactly aims at making maximum use of the potential of remote sensing data rather than aiming at calibration or validation of remote sensing data in order to come up with absolute values. Once this information has been compiled, it can extend the basis for sound water resources researches as well as constraining respective spatially distributed hydrological models right from the beginning.

The layout of the rest of the paper is as follows: Section 2 introduces the study area, data used, applied principal component analysis and statistical tests. Section 3 presents the results. Section 4 discusses the main findings. Finally, Section 5 offers an outlook on future studies.

2. Material and Methods

2.1. Study Area

The study region is the Wami River basin, located between 5°00′–7°27′ S and 36°00′–39°00′ E in east-central Tanzania (Figure 1). It has an area of approximately 41,170 km² and its elevation ranges from 0 to 2360 m a.s.l. (Figure 1). The river basin is separated into two major parts by the Eastern Arc Mountains (EAMs) which comprise the Rubeho, Ukaguru, Nguru, and Nguu mountain ranges (Figure 1). The geology comprises diverse lithologies derived from cratonic granitoids of the Precambrian age in the west, highly metamorphosed rocks of the Orogenic belts in the central and Neogene deposits in the eastern parts [10]. The river basin has been affected by faults, causing terrace and cascade flows at the western boundary of the coastal plain [10].
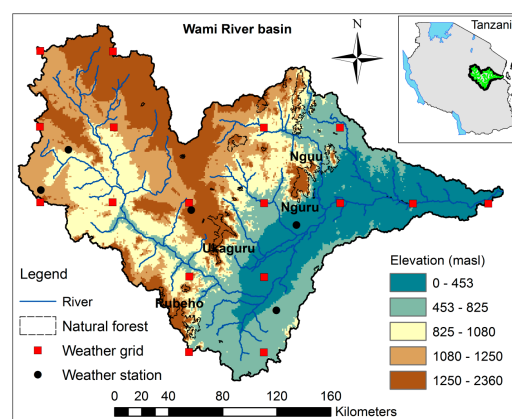

Figure 1. The Wami River basin (elevation, rainfall stations and grid points from Shuttle Radar Topography Mission (SRTM), Tanzania Meteorological Agency (TMA) and Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP) databases, respectively).

The topographic slopes are much steeper in the EAMs compared to other parts of the river basin and the slope angles range between 0° and 27° (Figure 2a). The predominant soil texture is loam and sandy-clay-loam which constitute 38% and 41% of the river basin area respectively (Figure 2b) [36]. The predominant land use classes are bushland (savannah), woodland (deciduous trees), grassland, ranch (savannah grassland), cropland (small-scale farming), irrigation areas, and natural forests (evergreen trees). Bushland, woodland and grassland cover 23%, 44%, and 20% of the basin area respectively [37] (Figure 2c). Ranch, cropland and irrigation areas cover about 10%. Natural forests which are predominantly located along the EAMs cover about 3%. The area of the EAMs is a globally important eco-region [38,39] and one of the world's hotspots of biological diversity [40,41]. Madoffe et al. [42] reported that in the year 1900 there was three times the amount of natural forest cover compared to the 2000s. In order to reduce further losses of biodiversity, logging was banned in the EAMs in the mid-1980s and 1990s and forest boundaries were restored in most reserves [42]. Other activities such as agriculture development, fuel wood collection, and charcoal burning were also prohibited in the protected areas. The river basin comprises the Saadani National park which is also very important for the downstream ecosystem.
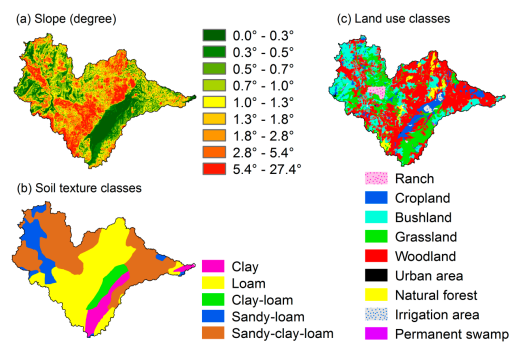
Figure 2. Topograhic slopes (derived from STRM-90m), soil texture classes of the year 2003 [36] and land use classes of the year 1997 [37].

The average daily air temperature in the river basin is between 21 and 27 °C (2000–2012), and the average rainfall is between 585 and 1175 mm per annum (2000–2012). The river basin has two major rainfall zones: A unimodal rainfall zone with one heavy rainfall season in the upstream area and a bimodal rainfall zone with two rainfall seasons in the downstream area. The unimodal rainfall occurs from late October to April (ONDJFMA). The bimodal rainfall comprises the light rainfall season between late October and December (OND) and the heavy rainfall season from March to May (MAM). The dry period starts in June and ends in September or early October. This is the period with no or little rainfall in the river basin from upstream to downstream. During this period the deciduous trees shed their leaves, most parts of the river flows are confined within the banks and most flood plains dry up. The average MODIS ET is between 353 and 1637 mm/year (2000–2013), with higher ET in the downstream area than in the upstream area (Figure 3).
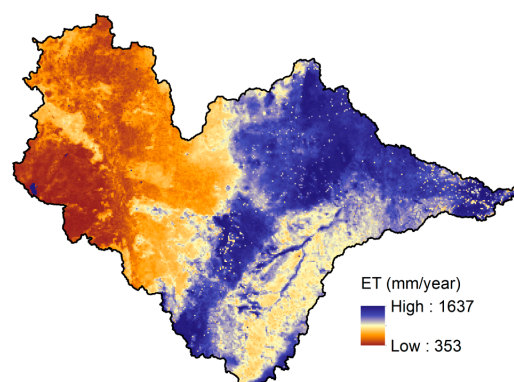

Figure 3. Average annual MODerate resolution Imaging Spectroradiometer evapotranspiration (MODIS ET) (2000–2013).

To some extent the hydrologic cycle of the Wami River basin is affected by domestic water supply, irrigation, recently introduced rainwater harvesting agriculture, increasing demand for charcoal, fuel wood, and timber [43,44]. The irrigation activities (e.g., sugarcane and rice plantations) accounted for an average of 96% of total abstracted water in the year 2010 [45,46].

2.2. Data

2.2.1. Land Surface and Meteorological Data

The land surface data for correlation and difference tests included elevation, slopes, land use, and soil texture data from various databases. A digital elevation model of 90 m resolution was downloaded from the Shuttle Radar Topography Mission (SRTM) database [47] and was used to classify topography and to compute slopes. Land use classes were derived from land use data

available at a scale of 1:250,000 from the Africover database [37]. The FAO [37] produced land use data from digitally enhanced high-resolution LANDSAT TM images acquired mainly in the year 1997. The soil texture classes of the year 2003 were derived from soil texture data available at a scale of 1:2,000,000 from the FAO-ISRIC [36] database.

The daily rainfall and air temperature data for climate variability tests were obtained from the Tanzania Meteorological Agency (TMA). To supplement the measured meteorological data, the reanalysed precipitation and air temperature data from the Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP, [48]) database were used.

Data for verifying the inferred hypotheses were depths to static water levels of wells and the visible satellite imageries. The depths to static water levels of groundwater wells in the Wami River basin were obtained from the Wami/Ruvu Basin Water Office [49]. The visible surface reflectance images (MOD09) of 10 June 2001 and 10 June 2012 both at a resolution of 500 m were downloaded from the United States Geological Survey (USGS) database (http://earthexplorer.usgs.gov/, accessed on 20 September 2016).

2.2.2. MODIS ET Data and Pre-Processing

ET data used were provided by MODIS. MODIS is an extensive program using the Terra and Aqua satellite data to provide a comprehensive series of global observations of the earth's land, oceans, and atmosphere in the visible and infrared regions of the electromagnetic spectrum [50]. The Terra earth observation system was launched in the year 1999 and the Aqua earth observation system in the year 2002 [50]. The Terra satellite goes from north to south across the equator and Aqua passes south to north over the equator [3].

The MODIS ET was calculated based on the Penman-Monteith equation [51]. The calculation was done on vegetated regions, but it did not include urban areas, permanent wetlands, and water bodies [51], thus these areas were also not considered in this study. The input data for the MODIS ET calculation were other MODIS products and meteorological reanalysis data. The MODIS products used were the global 1 km Collection 4 MODIS land cover type 2 (MOD12Q1), Collection 4 0.05-degree Climate Modeling Grid MODIS albedo (MOD43C1), global 1 km MODIS Collection 5 leaf area index and the fraction of photosynthetic absorbed radiation (MOD15A2) [51]. The global daily meteorological reanalysis dataset used were air temperature, air pressure, humidity, and solar radiation at a resolution of $1.00°$ by $1.25°$ from the NASA's Global Modeling and Assimilation Office [51]. Since the MODIS ET product is a composite index of a number of remote sensing products of various spatio-temporal resolutions and some global meteorological data, it was only used as a convenient index that has a hydrological interpretation. Thus, the focus of this study was not on accuracy of the MODIS ET rather on its spatio-temporal patterns.

For this study the most recent version (version 5: V005) of 8-day ET data with a spatial resolution of 1 km (MOD16A2) were used. The dataset was downloaded from the Numerical Terradynamic Simulation Group database of the research laboratory at the University of Montana in Missoula (http://www.ntsg.umt.edu/project/mod16, accessed on 17 October 2014) [51]. The downloaded MOD16A2 dataset contained 644 images between 1 January 2000 and 27 December 2013 (see Video S1 in the Supplementary Materials).

The 644 images of MOD16A2 with their corresponding quality control datasets (MOD16A2 QC) were extracted from HDF files and converted into the Geotiff format. Then the images were clipped to the size of the Wami River basin. The 644 MOD16A2 QC images were analysed for cloud cover. The average cloud cover per MOD16A2 image was 19%. Since clouds rarely repeated in the same areas, distortion due to clouds was negligibly small. Therefore, the computation of cloud masks was omitted as heavy fragmentation of the time series would occur if the masks were applied for even small clouds in every affected image and cumulatively applied to the entire time series [25]. In addition, MOD16A2 uses daily global reanalysis weather data as part of its input which potentially mitigates the

effect of missing data due to cloudiness [52]. For the purpose of consistency in the following sections, the MOD16A2 is referred to as the MODIS ET.

2.3. Analysis of Hydrological Behaviour

2.3.1. Principal Component Analysis

The primary objective of applying PCA to the MODIS ET dataset was to extract the most typical spatial patterns from the whole dataset of 644 images. The typical spatial patterns could include long-term prevailing patterns as well as peculiar patterns that were restricted to a single season or became clearly stronger or weaker in the long-term. Patterns of long-term changes usually point to possible effects of climate or land use change [29,53].

PCA, also known as the Karhunen-Loeve transformation or empirical orthogonal function [11,19], is a linear method that can be used to reduce the number of variables down to a few principal components (PCs) which explain most of the variance of the dataset [54–57]. The PCA approach basically decomposes a correlation matrix into eigenvectors and eigenvalues. The resulting eigenvectors, i.e., the principal components are un-correlated and are ordered by the fraction of variance they explain in descending order [58]. When PCA is applied to a set of time series, the first component usually describes the mean behaviour of the entire dataset and the following components reflect typical deviations from that mean behaviour [59].

Prior to the application of PCA, the 644 images of MODIS ET data covering the Wami River basin were flattened and arranged in matrix X (Equation (1)), where x represents a value of a pixel i from an image j. The variable n is the number of pixels in the image, thus each column in matrix X contains all pixels in a single image (48,132 pixels), whereas m is the number of images in the dataset (644 images).

$$
X = \begin{bmatrix}
x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1m} \\
x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2m} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{im} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{nm}
\end{bmatrix}
\tag{1}
$$

Each column in matrix X was normalized to a mean of zero and unit standard deviation using Equation (2) to obtain matrix Z (Equation (3)), where $\overline{x}_j$ and σ_j are the column mean and column standard deviation of matrix X, respectively. Normalization of the matrix prevented images with extreme conditions from dominating the PCA process.

$$
z_{ij} = \frac{x_{ij} - \overline{x}_j}{\sigma_j}
\tag{2}
$$

$$
Z = \begin{bmatrix}
z_{11} & z_{12} & \cdots & z_{1j} & \cdots & z_{1m} \\
z_{21} & z_{22} & \cdots & z_{2j} & \cdots & z_{2m} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
z_{i1} & z_{i2} & \cdots & z_{ij} & \cdots & z_{im} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
z_{n1} & z_{n2} & \cdots & z_{nj} & \cdots & z_{nm}
\end{bmatrix}
\tag{3}
$$

After that, the correlation matrix S (Equation (4)) containing the Pearson correlation coefficients between each column vector and others in matrix Z was calculated.

$$
S = corr\ (Z)
\tag{4}
$$

The Python routine *numpy.linalg.svd()* [60–62] was then used to perform the eigenvalue decomposition of the correlation matrix *S* which resulted in eigenvectors *U* and eigenvalues Λ, where *T* stands for transpose (Equation (5)). The eigenvalues are proportional to the fraction of variance explained by the respective eigenvectors [58].

$$S = U\Lambda U^T \tag{5}$$

The complete set of typical patterns was determined by multiplying the standardised dataset *Z* with the eigenvectors *U* to obtain the matrix of the principal component scores *P* (Equation (6)).

$$P = ZU \tag{6}$$

Each column vector of matrix *P* was then transformed to a 2-D pattern, inversely to the procedure described in Equation (1). Thus, any single principal component reflects only a fraction of the spatial pattern of a given MODIS ET image that has not been represented by any other principal component. These spatial patterns were then compared to those of elevation and slope, and also used to check for significant differences between different soil texture and land use classes in order to determine the land surface drivers of the hydrological behaviour in the river basin.

The extent to which a spatial pattern depicted by a single principal component plays a role for a single image of the MODIS ET dataset was quantified by the loadings [14]. Loadings *L* were calculated as the Pearson correlation coefficients of the pairwise comparisons of the scores of a single principal component from matrix *P* with standardized images from matrix *Z* (Equation (7)). Therefore, the spatial pattern of scores of a component was more pronounced on the highest absolute values of loadings, i.e., the extremes of loadings denoted periods where the respective spatial pattern was especially important.

$$L = corr\,(P,\, Z) \tag{7}$$

Because the 644 images of the MODIS ET constituted a time series, their respective loadings on principal components also constituted time series. Therefore, any single MODIS ET image can then be reconstructed by adding the scores of the PCs multiplied by their respective loadings.

The time series of loadings on selected PCs were analysed for clear seasonal patterns and long-term trends. For a more intuitive interpretation of the components, the scores and loadings on the first, third, and fifth PCs were reversed (multiplied by -1). The sign reversing led to the positive and negative scores and loadings representing high and low ET respectively. Subsequent analyses used these reversed PCs. To check for major land use changes, bands 1 (648 nm), 4 (555 nm), and 3 (470 nm) of the surface reflectance data were mapped as red, green, and blue (RGB) colours respectively to obtain the visible appearance of the Wami River basin in the years 2001 and 2012.

2.3.2. Trend Tests

The loadings on each PC constituted a time series of 644 points spanning between 1 January 2000 and 27 December 2013 at 8-day intervals. In order to check for temporal patterns of PCs, the loadings were tested for long-term trends using the pre-whitened Mann–Kendall test and Sen's slope [63–66]. The trend-free pre-whitening procedure was used to correct the time series of loadings for autocorrelation prior to the Mann-Kendall test [66]. The Mann–Kendall test was then used to look for a significant monotonic trend in a given dataset of loadings and Sen's slope was used to determine the magnitude of that trend, assuming that it was linear. For a *p* value less than or equal to 0.05 we rejected the null hypothesis of no long-term trend in the time series of loadings.

In order to investigate the effects of atmospheric forcing on the patterns of PCs, the meteorological effects test was also conducted. Prior to meteorological effects test, the daily rainfall and air temperature data spanning between the years 2000 and 2012 were converted into 8-day periods to yield the same temporal resolution as the loadings. The new time series of rainfall and air temperature at the stations

and grid points were then tested for long-term trends using the pre-whitened Mann-Kendall test and Sen's slope [63–66]. For a *p* value less than or equal to 0.05 we rejected the null hypothesis which stated that there was no long-term trend in the time series of rainfall or air temperature data.

2.4. Land Surface Drivers of Hydrological Behaviour

2.4.1. Statistical Dependence Test for Metric Data

The relationship between elevations or slopes and individual principal component scores were tested for significance using Kendall's rank correlation test [64]. This test is based on the ranks of data and uses the τ coefficient to measure the statistical dependence. Kendall's τ takes values between -1 and $+1$, where a positive τ value indicates a perfect positive relationship between two variables and a negative τ value an opposing relationship. For a *p* value less than or equal to 0.05 we rejected the null hypothesis which stated that Kendall's τ between principal component scores and elevations or slopes was equal to zero.

2.4.2. Statistical Difference Test for Nominal Data

The Kruskal-Wallis [67] and Wilcoxon [68] tests were used to check for significant differences of the scores of PCs for different soil texture and land use classes. Prior to the tests, all pixels of scores of a component in question were sorted and grouped in accordance with the soil texture or land use classes. Firstly, the Kruskal-Wallis test was used to detect differences in the scores between at least two of all soil texture or land use classes. When that was the case, the Wilcoxon post-hoc test was subsequently applied for pairwise comparison of scores of soil texture or land use classes. For both tests a *p* value less than or equal to 0.05 was used for rejection of the null hypothesis of no differences between different classes.

3. Results

3.1. Trends in Meteorological Data

The meteorological trend tests between the years 2000 and 2012 showed that, the *p* values for rainfall ranged between 0.31 and 0.97 (>0.05), and the *p* values for air temperature ranged between 0.08 and 0.56 (>0.05). Therefore, long-term trends of rainfall and air temperature were not significant at the 5% level. Thus, any trend found for individual PC necessarily had to be ascribed to different causes. However, these findings did not necessarily rule out the possibility that individual extreme meteorological events could have caused major effects on some ET patterns.

3.2. Principal Components

This study did not aim at a full understanding of the spatio-temporal patterns of ET. Instead, it aimed at extracting only the most dominant and peculiar information from 644 images of the MODIS ET dataset. Therefore, only PCs that reflected meaningful effects were considered. The study also followed the assumption that each single PC might reflect more than one effect or process that affect the hydrologic cycle of the river basin, partly only during very rare and extreme conditions. Thus, only the first five out of 644 PCs were selected based on the interpretability of their spatio-temporal patterns with respect to physical effects or processes. These first five PCs explained 81.2% of the variance in the dataset (Table 1).

Table 1. Eigenvalues, fractions of explained variance and cumulative proportions of the first five principal components (PCs).

Principal Component	Eigenvalue	Proportion of Explained Variance (%)	Cumulative Proportion (%)
1	405.9	63.0	63.0
2	55.8	8.7	71.7
3	36.5	5.7	77.4
4	13.5	2.1	79.5
5	10.9	1.7	81.2

Time periods on loadings were identified where the spatial pattern represented by the scores of an individual principal component played a major role (see Videos S2 to S6 in the Supplementary Materials). Correspondingly, a trend in the time series of loadings indicated whether the role of the respective spatial pattern increased or decreased in the long-term. Because data were normalized prior to the PCA, zero component scores represented average scores, positive scores denoted a higher than average markedness of the respective effect, and negative scores a lower than average markedness. However, what was considered an effect could equally be considered an inverse effect with a reversed sign (see Videos S3 and S4 in the Supplementary Materials). For example, a strong reduction in ET in the dry period compared to the mean pattern is necessarily inverse to a strong increase in ET in the wet period compared to the mean pattern.

The following sections present results of the five selected scores and loadings on PCs. The differences in the scores of PCs for different soil texture and land use classes are also presented. The long-term trends of loadings on these PCs are presented in Table 2. The results for relationships between scores and elevations or slopes are presented in Table 3.

Table 2. Trend analysis of the time series of PC loadings at 5% level of significance.

Principal Component	Mann-Kendall Trend Test (Two-Tailed p Value)	Sen's Slope (Change of Loading per Annum)
1	0.43	0.00
2	0.00	+0.01
3	0.00	+0.01
4	0.03	0.00
5	0.00	+0.02

Table 3. Kendall's rank correlation between the scores of the principal components and elevations or slopes.

Principal Component	Elevations		Slopes	
	τ	p Value	τ	p Value
1	−0.26	0.00	+0.16	0.00
2	+0.24	0.00	+0.04	0.00
3	+0.30	0.00	+0.24	0.00
4	−0.21	0.00	−0.13	0.00
5	−0.16	0.00	−0.11	0.00

3.2.1. First Principal Component

The first principal component (PC1) covered the largest fraction of explained variance (63.0%) (Table 1). The loadings on PC1 were usually close to 0.9 throughout the period, except for the first months of the year and did not show any significant trend (Figure 4a, Table 2). The map of average ET throughout the whole period and the scores of PC1 were nearly identical (cf. Figures 3 and 4b) with a coefficient of spatial correlation equal to 99.9%. Consequently, higher-order PCs depict typical deviations from PC1 during respective specific conditions.

The scores of PC1 divided the river basin into two parts which showed high and low ET in the downstream and upstream areas respectively (Figure 4b). In addition, natural forests exhibited relatively high scores indicative of systematically high ET (cf. Figures 2c and 4b). The Wami river riparian zone showed relatively high ET in the midstream and downstream areas, but it showed low ET in the upstream area (Figure 4b).

The scores of PC1 showed a significant negative correlation with elevations, and a significant positive correlation with slopes (Table 3). The scores of PC1 also exhibited significant differences among all soil texture classes, but loam and sandy-loam soils exhibited the highest and lowest medians respectively (Figure 4c). In addition, clay and sandy-clay-loam soils exhibited the narrowest and widest distributions of scores respectively. The scores between irrigation areas and woodland (*p* value = 0.15) as well as between bushland and ranch areas (*p* value = 0.14) did not differ significantly, but the remaining land use classes exhibited significant differences (Figure 4d). Natural forests exhibited the highest median, whereas bushland and ranch areas exhibited the lowest median of the scores of PC1. However, cropland and ranch areas exhibited the widest and narrowest distributions of scores respectively.
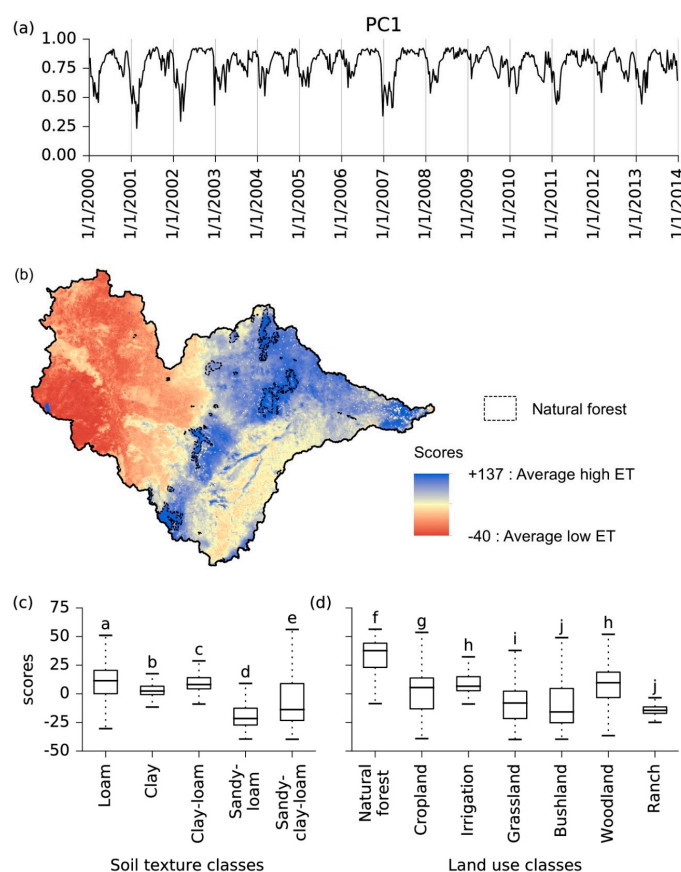

Figure 4. Time series of the loadings on the first principal component (PC1) (**a**). Map of the scores of PC1 and natural forest areas (**b**). Quartiles of the scores of PC1 for different soil texture (**c**) and land use (**d**) classes, similar score distributions at the 5% level of significance are marked with the same small letter.

3.2.2. Second Principal Component

The second principal component (PC2) covered 8.7% of explained variance (Table 1). The loadings on PC2 showed an increasing trend (Table 2), but there were also single extreme events in the years 2002, 2006, and 2007 (Figure 5a). The time series of loadings exhibited a clear seasonal pattern with the

highest loadings (about 0.6) at the end of the annual dry season in September and October (Figure 5a). Negative peaks of loadings occurred in March but did not undercut −0.4. This illustrates that the spatial pattern of PC2 predominantly captured the September–October period as compared to March.

Low scores, i.e., lower than average ET in the September-October period were found in most lowland areas in the downstream part of the river basin (cf. Figures 1 and 5b). Higher than average ET in this period were found in the loam soil region in the downstream part, and in the sandy-clay-loam soil region in the western and north-western parts of the river basin (cf. Figures 2b and 5b). In these regions the depth to static water levels was less or equal to 10 m below the ground (Figure 5b). By using the inverse proportionality, higher and lower than average ET in the September–October period might correspond to static water depths of less or equal to 10 m and greater or equal to 23 m below the ground respectively.

Natural forest areas clearly stand out as sharply delineated regions of high scores, i.e., higher than average ET in the September–October period (cf. Figures 2c and 5b). Groundwater data were not available for these regions in the mountainous region of the EAMs. However, a shallow depth to groundwater is not very likely here.
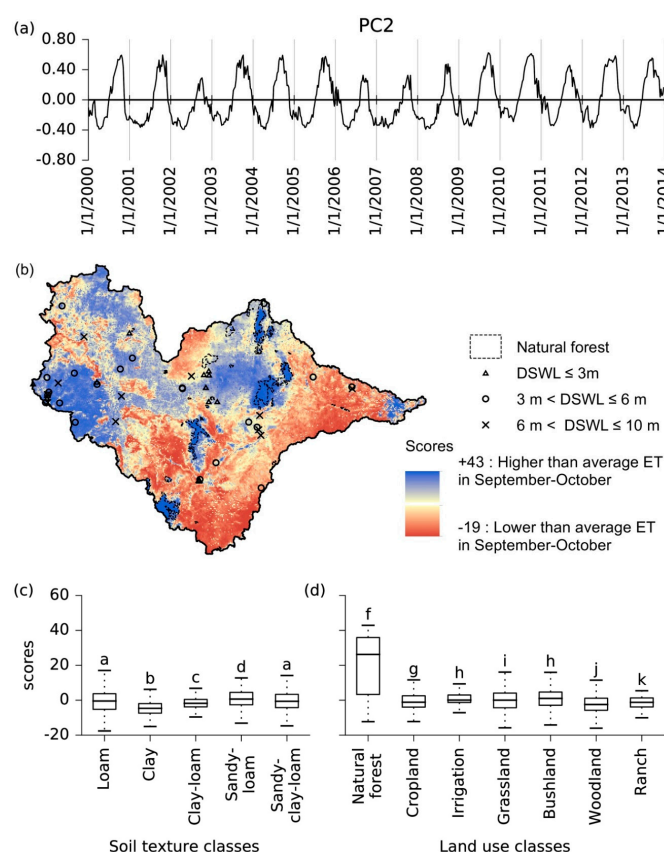

Figure 5. Time series of the loadings on the second principal component (PC2) (**a**). Map of the scores of PC2, natural forest areas and depths to static water levels (DSWL) measured below the ground (**b**). Quartiles of the scores of PC2 for different soil texture (**c**) and land use (**d**) classes, similar score distributions at the 5% level of significance are marked with the same small letter.

In the September–October period, the scores of PC2 significantly increased with both elevations and slopes, although the relationship was stronger for elevations than for slopes (Table 3). Unlike other soil texture classes, the scores of PC2 for sandy-clay-loam and loam soils did not differ significantly (Figure 5c). However, sandy-loam and clay soils exhibited the highest and lowest medians respectively. The narrowest and widest distributions of the scores were exhibited by clay-loam and loam soils

respectively. Among the land use classes, only the scores of PC2 for bushland and irrigation areas did not differ significantly. Moreover, natural forests and woodland exhibited exceptionally the highest and lowest medians respectively (Figure 5d). Nevertheless, natural forest and irrigation areas exhibited the widest and narrowest distributions of scores of PC2 respectively.

3.2.3. Third Principal Component

The third principal component (PC3) explained 5.7% of the variance (Table 1). Figure 6a shows that the loadings on PC3 exhibited a clear seasonal pattern with highly positive peaks (about 0.6) in the January–February period and minor peaks in the August–September period. Moderate negative peaks (about −0.4) occurred in the May–June period and in the October–November period. The loadings showed a significant increase and individual strong events in the January–February periods in the years 2010, 2011, and 2013 (Figure 6a, Table 2).

High scores of PC3 indicated higher than average ET in the upstream area of the river basin during the January–February period (Figure 6b). In contrast, ET was relatively low in the downstream area and in the ranch region in the upstream area during this period (cf. Figures 2c and 6b). Similarly as for the previous PCs natural forests exhibited relatively high ET during this period.

The scores of PC3 correlated equally well positively with both elevations and slopes (Table 3). There were significant differences between all soil texture classes, but sandy-loam and clay-loam soils exhibited the highest and lowest medians respectively (Figure 6c). However, sandy-loam and clay have the widest and narrowest distributions of scores respectively. The scores of PC3 for land use classes showed that, bushland and woodland were similar but the remaining classes differed significantly. The highest and lowest medians were exhibited by natural forest and irrigation areas (Figure 6d). Nevertheless, bushland and woodland exhibited the widest distributions, whereas irrigation areas exhibited the narrowest distribution.
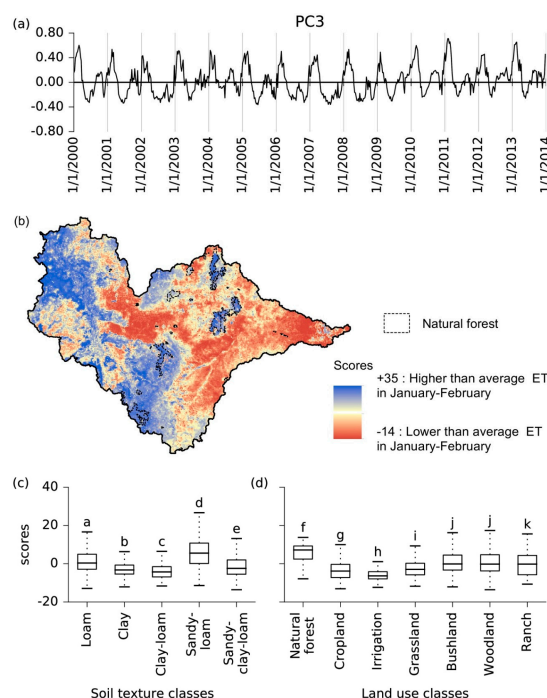

Figure 6. Time series of the loadings on the third principal component (PC3) (**a**). Map of the scores of PC3 and natural forest areas (**b**). Quartiles of the scores of PC3 for different soil texture (**c**) and land use (**d**) classes, similar score distributions at the 5% level of significance are marked with the same small letter.

3.2.4. Fourth Principal Component

The fourth principal component (PC4) covered 2.1% of the variance (Table 1). In contrast to the first three principal components, the loadings on PC4 exhibited only weak seasonal patterns with maximum values between October and February, but with pronounced short-term fluctuations (Figure 7a). The loadings exhibited a significant increase with extreme positive values in the end of January 2007 and extreme negative values in March 2013 (Figure 7a, Table 2).

The scores of PC4 showed a remarkably spatial pattern, clearly independent of the topography (cf. Figures 1, 2a and 7b). The scores divided the river basin into three nearly homogeneous parts with high ET in the upstream and downstream areas, and low ET in the midstream area, west of the mountainous regions.

The scores of PC4 negatively correlated with both elevations and slopes (Table 3). However, the relationship was stronger for elevations than for slopes. The scores of PC4 also showed that all soil texture classes exhibited different distributions significantly. Sandy-loam and loam soils exhibited the highest and lowest medians respectively (Figure 7c). However, the widest distribution of scores was exhibited by sandy-clay-loam, whereas the narrowest distribution was exhibited by clay-loam. For land use, all classes were significantly different except for bushland and cropland. The land use classes which exhibited the highest and lowest medians of the scores of PC4 were bushland and ranch areas respectively (Figure 7d). Irrigation areas exhibited the narrowest distribution of scores, but woodland and bushland exhibited the widest distributions of scores (Figure 7d).
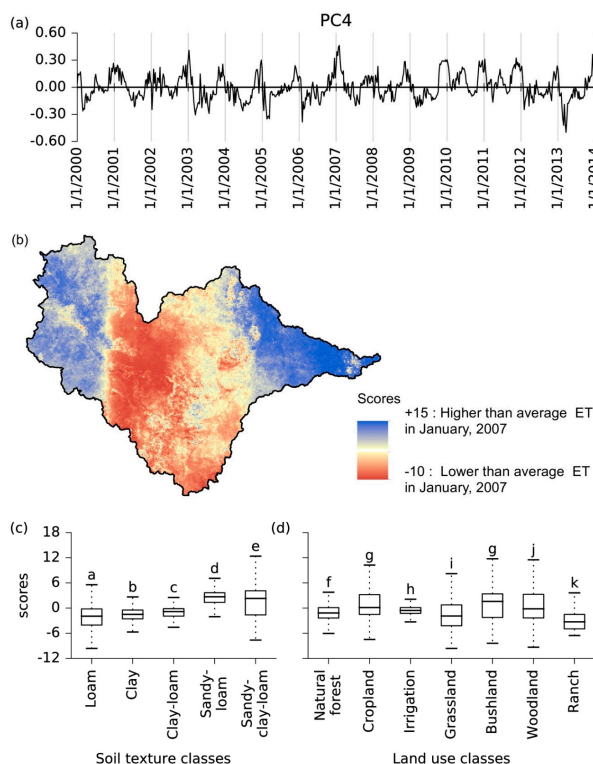

Figure 7. Time series of the loadings on the fourth principal component (PC4) (**a**). Map of the scores of PC4 (**b**). Quartiles of the scores of PC4 for different soil texture (**c**) and land use (**d**) classes, similar score distributions at the 5% level of significance are marked with the same small letter.

3.2.5. Fifth Principal Component

The fifth principal component (PC5) covered 1.7% of the variance (Table 1). The time series of loadings on PC5 exhibited substantial short-term variation and lacked a clear seasonal pattern

(Figure 8a). However, they exhibited the clearest long-term trend among the first five components with an average increase of 2% per annum (Table 2).

High scores were found in the downstream parts of the river basin and in the lowlands east and south of the EAMs (cf. Figures 1 and 8b). In contrast, consistently low scores occurred especially in the region north and west of the EAMs, except for regions of natural forests that exhibited relatively high scores irrespective of their location in the river basin (cf. Figures 2c and 8b).

Both elevations and slopes were significantly correlated with the scores of PC5 (Table 3). We also found significant differences among the scores of PC5 for all soil texture classes, but the highest and lowest medians were exhibited by clay and loam soils respectively (Figure 8c). However, loam soil also has the widest distribution of scores, whereas clay has the narrowest distribution of scores. For land use, classes with similar patterns of scores were cropland and grassland, but other classes were significantly different (Figure 8d). The highest and lowest medians of the scores of PC5 for land use classes were exhibited by natural forest and ranch areas respectively. But the widest distributions of scores were exhibited by natural forest and woodland, whereas the narrowest distributions were found in irrigation and ranch areas.

Figure 9 shows the land covers in the Wami River basin in the years 2001 and 2012. The images show that in the period between the years 2001 and 2012, forests diminished in the north of the river basin and along the EAMs (cf. Figures 1, 2c, 8b and 9). However, natural forests appeared to be the same during that period (cf. Figures 8b and 9). Moreover, slight changes of land cover in the west of the river basin were also found.
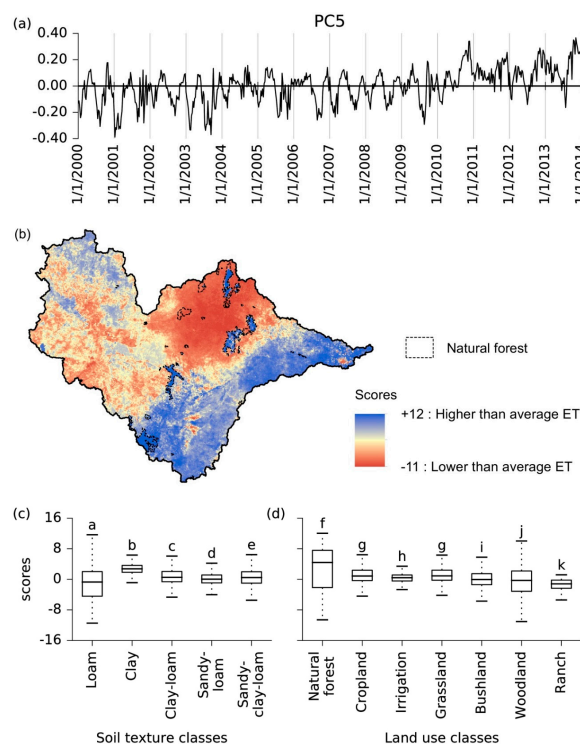

Figure 8. Time series of the loadings on the fifth principal component (PC5) (**a**). Map of the scores of PC5 and natural forest areas (**b**). Quartiles of the scores of PC5 for different soil texture (**c**) and land use (**d**) classes, similar score distributions at the 5% level of significance are marked with the same small letter.
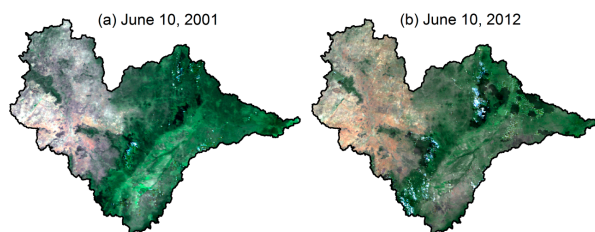
Figure 9. Satellite imageries for the years 2001 (**a**) and 2012 (**b**) mapped as red, green, and blue (RGB) (648, 555, and 470 nm). The deep and light green colours represent dense and sparse forests respectively, other colours represent scattered vegetation (not discernible at 500 m resolution).

4. Discussion

4.1. General Approach

The objective of this study was to infer information about hydrological behaviour in the Wami River basin as a basis for sustainable water resources and land use management as well as constraining subsequent hydrological models in a data-scarce environment. Thus, contrary to numerous other applications of the PCA approach, here the basic idea was not to assign processes or specific variables of influence to the PCs, but to identify prevailing and peculiar spatio-temporal patterns that were restricted to certain specific boundary conditions.

Correspondingly, this study focused on the spatial patterns of ET and how these changed during specific boundary conditions rather than aimed at a quantitative assessment of water fluxes. The latter is known to be prone to substantial uncertainties and to require substantial efforts of ground truth measurements. In contrast, the approach followed in this study makes maximum use of the potential of remote sensing data, i.e., grasping the spatial patterns rather than giving absolute numbers. However, there are still some sources of uncertainty that need to be considered.

The extent of meaningful physical information extracted from a PCA of remotely sensed ET can vary with the resolution (i.e., <1 km) and the type of remotely sensed data used. Yang et al. [69] illustrated that different remote sensing ET products use different formulae and assumptions, that result in different estimates of ET. Uncertainties in remote sensing ET estimates also depend on geographic location that affects climate (i.e., cloud cover, solar radiation intensity, etc.). Since this study was conducted in the equatorial region in a tropical climate (i.e., high solar radiation intensity, etc.) and focused on relative spatial patterns rather than on absolute ET values, the uncertainties of the MODIS ET were considered minimal. In addition, systematic errors of remote sensing data have little impact on long-term anomalies [5].

Neither rainfall nor air temperature exhibited a significant trend during the study period. Consequently, meteorological effects between the years 2000 and 2012 are considered to be negligible. Thus, significant trends in the scores of individual PCs offer some evidence for systematic shifts in land cover. The following sections discuss the inferences on hydrological behaviour and possible land surface drivers.

4.2. First Principal Component: Mean Behaviour of ET

All 644 images were strongly positively correlated with PC1 which did not change in the long-term. In fact, PC1 often expresses the most cumulative information from the dataset [30]. The similarity between the map of average ET and the scores of PC1 suggest that PC1 mainly reflects the mean behaviour of ET in the river basin. Similarly, PCA of time series often yields a first component that depicts the mean behaviour [17,18,59,70]. This component explained 63.0% of the variance of spatial and temporal patterns of ET, thus it represents the most prevailing conditions in the river basin. The respective spatial pattern of ET likely reflects the spatial patterns of vegetation type and density

in the river basin. For example, the downstream part of the river basin exhibited higher ET than the upstream part. In fact, Figure 9 attests higher canopy density in the former than in the latter.

The negative correlation for the scores of PC1 against elevations indicated higher water availability for ET in low elevations than in high elevations. This is due to higher rainfall in low elevations in the downstream area than in high elevations in the upstream area. It is because the EAMs prevent the influence of Indian Ocean cyclonic rains from reaching the upstream area. Hence, heavier rainfall is encountered in the downstream area than in the upstream area [45,71]. The significant, but small positive correlation for the scores of PC1 against slopes was caused by the relatively high ET in the natural forest along the EAMs, where steeper slopes are found.

All soil texture classes differ significantly with respect to the scores of PC1. This could partly be ascribed directly to the different water retention capacities of different soils. On the other hand, land cover and correspondingly vegetation density are also related to soil texture, which has an indirect effect on ET. Soil texture classes are not randomly distributed within the catchment, thus different ET from different classes might partly be related to the spatial pattern of rainfall as well. The similarity of ET between the irrigation areas and woodland suggests that due to irrigation activities, the rice and sugar cane plantations evaporate and transpire as much water as woodland which has deeper roots. Correspondingly, the similarity between bushland and the ranch region could be due to the fact that bushes remain at the ranch even after grazing and may transpire similarly to bushland. The widest distributions of ET in loam, sandy-clay-loam, cropland, and bushland areas were caused by scattering of these soil texture and land use classes across the river basin, thus they exhibited both low and high ET in the upstream and downstream areas respectively (Figures 2 and 4). However, the narrowest distributions of ET in clay, clay-loam, irrigation and ranch areas were caused by localized nature of these soil texture and land use classes, thus they exhibited localized high and low ET in the upstream and downstream areas respectively.

4.3. Second Principal Component: Dry Season Effects

The clear seasonal patterns shown by the loadings on PC2 in September–October and March represented the end of the dry period and the middle of the major rainy season (wet period) respectively. In contrast, the significant trend of the loadings on PC2 seemed primarily to be due to single, extreme dry events in the years 2002, 2006, and 2007 which led to relatively low ET across the river basin. Thus, we assume that the trend did not really reflect a continuous increase over time.

The spatial pattern of the scores of PC2 exhibited enhanced ET at the end of the dry period in the western part, north-western part and further downstream. In these regions the depths to static water levels in the wells were less or equal to 10 m below the ground [10,49]. The most plausible reason is that, plants have access to shallow groundwater that is related to geological structures in these regions. In the western part of the river basin, granites and migmatites are abundant, where the graben and horst formed by geological faults striking NNE-SSW are found [10]. In this area, groundwater is mainly restricted to the weathered and fractured part. The faults increase the water-holding capacity of granitic aquifers. In the north-western part, the graben and horst formed by NNE-SSE-striking faults act as the main groundwater reservoir [10]. Further downstream, composite metamorphic crust domains and granulites, gneisses and migmatites are found, but the geological fault is not well developed. Correspondingly, the aquifers are not as productive as in the western part of the river basin [10].

Apart from this, clearly higher than average ET that was detected at the end of the dry season in natural forest areas at high elevations, cannot be related to shallow groundwater. These areas clearly stood out compared to all other vegetation types, especially with respect to exceptionally sharp boundaries to adjacent areas with different vegetation cover. It is remarkable that this holds for PC1, PC3, and PC5 as well. In the case of PC2, that has extended ET even at the end of the dry season, it is due to canopies of evergreen forest intercepting nightly fog [72]. Trees in these cloud forests do not shed leaves even in the dry period.

Usually, during the dry season most of the forest plants shed their leaves due to the decrease of soil moisture. However, in the upstream area, vegetation in the areas with water table close to the surface exhibited relatively high ET as compared to the downstream area. That explains the significant positive correlations of scores of PC2 against both elevations and slopes.

Significantly differing scores for different soil texture classes except for loam and sandy-clay-loam give some evidence for different water retention properties that play a role especially during the dry season. The similar ET between loam and sandy-clay-loam is due to shallow groundwater that is largely found in the upstream area where the two soil texture classes exist. No significant differences in ET were found between irrigation areas and bushland. Both land cover classes exhibited low ET at the end of the dry period. This might be the case just by coincidence. The bushland ET was presumably restricted by low water availability, whereas in the irrigation areas this is the crop harvesting period (end of June–end of September). The widest distributions of ET in loam, sandy-clay-loam and natural forest areas were caused by the north-south spread of loam and natural forest as well as the presence of sandy-clay-loam in the upstream and downstream areas, thus they exhibited both high and low ET respectively (Figures 2 and 5). In contrast, the narrowest distributions of ET in clay and ranch areas were caused by their localized presence in the downstream and upstream areas respectively.

4.4. Third Principal Component: General Spatial Patterns of Rainfall

The highest loadings on PC3 in the January–February, May–June, and October–November periods coincided with the peaks of the ONDJFMA, MAM, and OND rainfall seasons. Thus, the loadings on PC3 reflected primarily ET effects due to rainfall distribution in the river basin. The spatial pattern of PC3 shows that the rainy period of January–February caused higher ET in the upstream area due to the unimodal (ONDJFMA) rainfall whereas in the downstream areas there was no rainfall in that period. As a result, the scores of PC3 have positive correlations against both elevations and slopes in the January–February period. The observed significant trend of the loadings on PC3 was caused by relatively heavier January–February rainfall events in the years 2010, 2011, and 2013 in the upstream area. Thus, we assume that the trend did not really reflect a continuous increase. In the ranch region low ET might be due to late vegetation sprout. The natural forests in the downstream area exhibited relatively high ET in the January–February period when there were no heavy rainfall events because the plants in the EAMs are cloud forests [72].

The scores of PC3 for different soil texture classes differed significantly. Despite the shallow groundwater (see Section 4.3) and rainfall effects during the peak of ONDJFMA, larger water holding capacity of loam soil in the upstream area contributed to the availability of more water than the sandy soil in the downstream area. In addition, different vegetation cover played different roles. The similarity of scores of PC3 between bushland and woodland in the January–February period may be caused by vegetation sprout of the deciduous trees which shed their leaves in the June–September/early October period in the woodland. Thus, during the sprout period the rate of transpiration from bushland and woodland may not be very different.

4.5. Fourth Principal Component: Single Major Rainstorms

The spatial pattern of PC4 differs remarkably from the other four PCs because it is not clearly related to any topographical structure or vegetation pattern. The transition zone between positive and negative scores west of the EAMs, stretching in the nearly perfect north-south direction (Figure 7b), provides strong evidence for atmospheric forcing. In addition, the lowest values encountered west of the EAMs, but not in the further upstream areas reflected a lee effect with respect to moisture transport from the east and the west. The time series of loadings on this component exhibited less clear seasonal patterns compared to those of PC3. High loadings on PC4 were restricted to single, relatively short periods, especially at the end of January 2007 and in March 2013 (Figure 7a). Thus, this component seems to reflect deviations from the usual spatial pattern of rainfall and subsequent ET during single

periods. Since the focus of the study is on the prevailing effects or processes and not on single extreme events, we will discuss the event at the end of January 2007 only briefly, as a major extreme event.

At the end of January 2007 the relatively high ET in the upstream and downstream areas is ascribed to the preceding heavy rainfall in OND, 2006. The OND, 2006 rainfall caused flooding in most of eastern Tanzania. It was related to exceptionally high sea surface temperature in the western Indian Ocean and Indo-Pacific El Nino effects [73]. That was the largest extreme rainfall event since OND, 1997. The relatively lower ET in the midstream area could be caused by the lee effect of the EAMs with respect to strong fluxes of humid easterlies and weak fluxes of low level humid westerlies from the southern Congo basin [71,74], that differs due to different air humidity transport velocities. The negative correlation for the scores of PC4 against elevations and slopes might be caused by relatively low ET in the midstream region, where elevations and slopes are the highest in the EAMs.

4.6. Fifth Principal Component: Long-Term Trend of Land Use

Although PC5 explained less variance than any of the first four PCs, the loadings on PC5 exhibited the strongest long-term trend, including a distinctive, almost stepwise increase in 2010. Because neither rainfall nor temperature exhibited corresponding trends, land use change was considered the most probable cause. In fact, often components of lower rank order are associated with long-term changes [29,53]. Combining the trend of the loadings with the spatial pattern of the component scores, the results pointed to deteriorating land cover in the west and north of the EAMs. The visible satellite imageries also evinced diminishing forest cover in the north of the river basin and along the EAMs as compared to other areas. Positive scores further downstream, in the lowlands east of the EAMs and in the south of the river basin can be interpreted as reflecting areas that did not exhibit any trend at all. This part might have been intensively used already before the start of the MODIS ET dataset.

Schaafsma et al. [41] found that commercial charcoal production is still practiced in the lower woodland areas of the EAMs (with some production centres within the boundaries of the EAM blocks), thus causing degradation of woodland in the EAMs. FBD [75] also reported that the EAMs exhibit rapid land cover change, having lost 11% of their primary forests and 41% of their woodland vegetation since 1975. This conversion is driven by clearance for farmland, as well as by increasing demand for timber and fuel wood [40]. This supports the argument that land cover change was caused by the deforestation of woodland vegetation.

Despite being in the same area, some natural forests did not exhibit deteriorating land cover because they are within the conservation areas [42,76,77], thus they were not affected by anthropogenic activities. Currently, the vast majority of natural forests in the EAMs are under different forms of legal protection, with most falling within the category of "National Forest Reserves" managed for protection of water resources, soil erosion prevention, and biodiversity conservation [78–80]. However, Tabor et al. [81] argued that forest loss was still happening in the protected areas, though to a far lower extent than outside protected areas.

Correlation of scores of PC5 with elevations and slopes were significant, but rather weak, thus will not be discussed in more detail. The similarity of scores of PC5 between cropland and grassland may be caused by the growth of grass in areas that were initially cultivated, but later abandoned by local farmers for various reasons including loss of soil fertility after being used for some time.

4.7. Implications for Water Resources Management and Distributed Hydrological Models

A lot of information is required for setting up and calibrating distributed hydrological models. Although some of the calibration parameters are closely related to land surface conditions, the spatial variability of these parameters in a large river basin such as the Wami River basin is very difficult to assess. We hypothesized that much can be learned about hydrological behaviour and hydrologically relevant structures by analysing time series of ET data determined by remote sensing. Findings gained using this approach are summarized in the following paragraphs.

About 63% of the spatio-temporal variance of the ET data was already captured by the mean spatial pattern depicted by PC1. It significantly depended on elevations and slopes. The widest distributions of ET shown by PC1 in loam, sandy-clay-loam, cropland, and bushland areas suggest that the roles of these soil texture and land use classes are very crucial in determining the mean behaviour of ET in the river basin. In contrast, the narrowest distributions of ET in clay, clay-loam, irrigation, and ranch areas demonstrate their little effects on the mean behaviour of ET in the river basin. However, PC1 also informed that the effects of different soil textures are significantly different, thus all soil texture classes are absolutely crucial for the hydrological model of the river basin. Likewise, most of the different land use classes are significantly different, except for woodland and irrigation areas, and bushland and ranch areas. Therefore, from the perspective of modelling the mean behaviour of ET, the land use classes of irrigation and ranch areas can be replaced with woodland and bushland respectively, without any loss of information.

However, PC1 did not explicitly capture the effects of irrigation. This was not consistent with our expectations prior to the analysis. It cannot be explained by insufficient spatial resolution of the ET data. In fact, irrigation areas are known to extend more than 15 km^2 in the eastern part of the river basin, thus they should have been discernible. In contrast, it has to be concluded that the current irrigation density is relatively low and does not affect the hydrologic cycle in the river basin to the extent we presumed. Similar to PC1, the irrigation areas did not stand out in the spatial pattern depicted by PC2. This is presumably due to the fact that harvest occurs in the dry season. Thus, there is hardly any need for extended irrigation during this season. However, information provided by PC2 helped to better understand the reason for enhanced ET in other parts of the river basin. Sustained high ET in the western and north-western parts at the end of the dry season pointed to shallow groundwater that would be available for plant root uptake. Conducting sufficient spatially distributed measurements of the shallow groundwater or aquifer in the large river basins is not economically feasible. Therefore, the shallow aquifer boundary condition from PC2 can be used to emulate spatially distributed critical depths of water in shallow aquifer routines in hydrological models. In many distributed hydrological models this information is used to control the movement of water from the aquifers to the root zones.

In contrast, shallow groundwater cannot explain the outstanding high ET of natural forests with clear-cut boundaries to adjacent land cover classes. This feature was dominant for PC1, PC2, PC3, and PC5. It sheds some light on the effects of land cover changes around the natural forest areas. In terms of ET, even intensively cultivated areas cannot compensate for high water availability of natural forests as long as the former is not irrigated. Apparently this is due to interception of nightly fog in the canopies of evergreen trees in the natural forest areas, thus ensuring high water availability even during extended dry periods [72]. Thus, any conversion or impairment of the natural forests obviously would have major effects on the hydrologic cycle in the river basin. The most important difference between natural forests and other land cover classes seems to be their ability to maintain high physiological activity even under dry conditions, pointing to better adaptation to local climatic conditions. On the other hand, PC5 provided some evidence that the protection of the natural forests seems to have been highly effective during recent years. However, close to the nature reserve regions, land use change seemed to have had a major impact, especially in the northern part of the river basin close to the EAMs. This land cover change, especially the step shift around 2010, should be considered both for water resources management and in spatially distributed hydrological models of the Wami River basin.

The loadings on PC3 clearly reflected the unimodal seasonal rainfall in the upstream area and the bimodal rainfall in the downstream area. The spatial pattern of ET depicted by PC3 shows that the effects of January–Februry rainfall during ONDJFMA is substantial in the upstream areas as compared to the downstream areas. This is the minimum spatial differentiation of rainfall that a distributed hydrological model in the Wami River basin should account for. The pattern of PC4 reflected the lee effect during strong easterly rainfall periods and the weak westerly rainfall effects from the southern Congo basin [71,74], however, the pattern is more pronounced during extreme rainfall events.

It is remarkable that remote sensing ET data (i.e., PCs) gave information about prevailing and peculiar patterns of rainfall as well. In the short-term, ET was likely to decrease during single rainstorms due to high air humidity and low irradiation and temperature. However, given the limited temporal resolution of the MODIS ET data, this effect seemed to be negligible. In contrast, in the mid-term after heavy rainfall more water was available for plant uptake and thus, likely boosted ET, especially in regions with antecedent limited water supply. This plant response might be delayed due to the time needed for budding of perennial plants or for germinating and growth of annual plants, thus depending on the vegetation type and thus reflecting its spatial patterns. The summary of findings of principal components and their inferences for water resources management and distributed hydrological modelling in the Wami River basin is shown in Table 4.

Table 4. Findings and inferences for water resources management and distributed hydrological modelling.

PC	Main Features	Inferences for Hydrological Behaviour
1	Spatial pattern of long-term average ET (mean behaviour of ET).	Clear dichotomy between the upstream (low evapotranspiration (ET)) and downstream (high ET) parts of the river basin, partly due to a heavier March-May (MAM) rainy season in the latter. ET was exceptionally high in natural forests and loam soil, and very low in bushland and sandy-loam soil. No significant differences between ET of bushland and ranch areas. Irrigation of rice and sugar cane plantations obviously resulted in ET as high as in woodland. Loam, sandy-clay-loam, bushland and cropland areas have widespread effects on average ET across the river basin. Clay, clay-loam, current irrigation and ranch areas have localized effects on average ET in the river basin.
2	Regions of extended high ET at the end of the dry season.	Regions of shallow groundwater, accessible by plant roots in the dry season. Outstanding role of fog interception in regions of natural cloud forests. Effect of irrigation not visible during the dry season due to earlier harvest. No significant differences between loam and sandy-clay-loam during the dry season. High importance of this dry season pattern in the June–September periods in the years 2002, 2006, and 2007.
3	Spatial effect of rainfall seasons.	Unimodal (October-April (ONDJFMA)) and bimodal (October-December (OND) and MAM) rainfall distributions in the upstream and downstream parts of the river basin respectively. ONDJFMA rainfall during the January–February periods increases ET in the upstream part of the river basin, at high elevations and steep slopes.
4	Lee effect of strong humid easterlies and effects of weak humid westerlies.	Effect on the spatial pattern of ET in the river basin due to strong rainfall from the east and weak rainfall from the west of the Eastern Arc Mountains (EAMs).
5	Long-term change of land use.	Long-term and spatially almost homogeneous reduction of ET due to massive deforestation of woodland vegetation northwest of the EAMs, except for the forest nature reserves.

5. Conclusions

Planning and managing water resources, and modelling to support the former requires a sound understanding of the hydrologic cycle of the river basin. However, many regions of the world lack extensive monitoring networks of river gauges and groundwater observation wells required for hydrologists to base their advice and their models upon. On the other hand, high-quality remote sensing products are now globally available. We recommend using these products for water resources management and planning, for hydrological system analysis prior to setting up the models and testing them. In this study, inferences regarding hydrological behaviour were derived from a PCA of MODIS ET data.

We showed how a time series of MODIS ET data could be used to assess hydrological behaviour in the data-scarce region of the Wami River basin in Sub-Saharan Africa. A PCA was applied to extract prevailing and peculiar spatial patterns for different seasons and possible long-term trends. The main results elucidated the mean behaviour of ET, illustrated the distinctive behaviour of natural forests, helped to identify regions of shallow groundwater, and clearly pointed to long-term shifts in degrading vegetation cover in parts of the river basin. In addition, the results allowed reduction of the number of land use classes to be considered in the distributed hydrological model of the river basin.

We conclude that these findings will highly improve the task of setting up and calibrating or validating the subsequent distributed hydrological model in the Wami River basin.

References

1. Li, Z.-L.; Tang, R.; Wan, Z.; Bi, Y.; Zhou, C.; Tang, B.; Yan, G.; Zhang, X. A review of current methodologies for regional evapotranspiration estimation from remotely sensed data. *Sensors* **2009**, *9*, 3801–3853. [CrossRef] [PubMed]

2. Eswar, R.; Sekhar, M.; Bhattacharya, B.K. A simple model for spatial disaggregation of evaporative fraction: Comparative study with thermal sharpened land surface temperature data over India. *J. Geophys. Res. Atmos.* **2013**, *118*, 12029–12044. [CrossRef]

3. Kiptala, J.K.; Mohamed, Y.; Mul, M.L.; Van Der Zaag, P. Mapping evapotranspiration trends using MODIS and SEBAL model in a data scarce and heterogeneous landscape in Eastern Africa. *Water Resour. Res.* **2013**, *49*, 8495–8510. [CrossRef]

4. Jovanovic, N.; Mu, Q.; Bugan, R.D.H.; Zhao, M. Dynamics of MODIS evapotranspiration in South Africa. *Water SA* **2015**, *41*, 79–90. [CrossRef]

5. Alemu, H.; Senay, G.B.; Kaptue, A.T.; Kovalskyy, V. Evapotranspiration variability and its association with vegetation dynamics in the Nile Basin, 2002–2011. *Remote Sens.* **2014**, *6*, 5885–5908. [CrossRef]

6. Munch, Z.; Conrad, J.E.; Gibson, L.A.; Palmer, A.R.; Hughes, D. Satellite earth observation as a tool to conceptualize hydrogeological fluxes in the Sandveld, South Africa. *Hydrogeol. J.* **2013**, *21*, 1053–1070. [CrossRef]

7. Ramoelo, A.; Majozi, N.; Mathieu, R.; Jovanovic, N.; Nickless, A.; Dzikiti, S. Validation of global evapotranspiration product (MOD16) using flux tower data in the African Savanna, South Africa. *Remote Sens.* **2014**, *6*, 7406–7423. [CrossRef]

8. Yilmaz, M.T.; Anderson, M.C.; Zaitchik, B.; Hain, C.R.; Crow, W.T.; Ozdogan, M.; Chun, J.A.; Evans, J. Comparison of prognostic and diagnostic surface flux modeling approaches over the Nile River basin. *Water Resour. Res.* **2014**, *50*, 386–408. [CrossRef]

9. Deus, D.; Gloaguen, R. Remote sensing analysis of lake dynamics in semi-arid regions: Implication for Water Resource Management. Lake Manyara, East African Rift, Northern Tanzania. *Water* **2013**, *5*, 698–727. [CrossRef]

10. The Global Water for Sustainability Program-Florida International University (GLOWS-FIU). *Water Atlas of Wami/Ruvu Basin, Tanzania*; ISBN 978-1-941993-01-9; Florida International University: Miami, FL, USA, 2014.

11. Rezaei, A.; Sayadi, M.H. Long-term evolution of the composition of surface water from the River Gharasoo, Iran: A case study using multivariate statistical techniques. *Environ. Geochem. Health* **2015**, *37*, 251–261. [CrossRef] [PubMed]

12. Sayyed, G.; Sayadi, M. Variations in the heavy metal accumulations within the surface soils from the Chitgar industrial area of Tehran. *Proc. Int. Acad. Ecol. Environ. Sci.* **2011**, *1*, 36–46.

13. Petersen, W.; Bertino, L.; Callies, U.; Zorita, E. Process identification by principal component analysis of river water-quality data. *Ecol. Model.* **2001**, *138*, 193–213. [CrossRef]

14. Gao, B.-B.; Wang, J.-F.; Fan, H.-M.; Xu, K.; Hu, M.-G.; Chen, Z.-Y. A stratified optimization method for a multivariate marine environmental monitoring network in the Yangtze River estuary and its adjacent sea. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 1332–1349. [CrossRef]

15. Gao, H.; Lv, C.; Song, Y.; Zhang, Y.; Zheng, L.; Wen, Y.; Peng, J.; Yu, H. Chemometrics data of water quality and environmental heterogeneity analysis in Pu River, China. *Environ. Earth Sci.* **2015**, *73*, 5119–5129. [CrossRef]

16. Hohenbrink, T.L.; Lischeid, G. Does textural heterogeneity matter? Quantifying transformation of hydrological signals in soils. *J. Hydrol.* **2015**, *523*, 725–738. [CrossRef]

17. Lehr, C.; Pöschke, F.; Lewandowski, J.; Lischeid, G. A novel method to evaluate the effect of a stream restoration on the spatial pattern of hydraulic connection of stream and groundwater. *J. Hydrol.* **2015**, *527*, 394–401. [CrossRef]

18. Lischeid, G.; Natkhin, M.; Steidl, J.; Dietrich, O.; Dannowski, R.; Merz, C. Assessing coupling between lakes and layered aquifers in a complex Pleistocene landscape based on water level dynamics. *Adv. Water Resour.* **2010**, *33*, 1331–1339. [CrossRef]

19. Thomas, B.; Lischeid, G.; Steidl, J.; Dannowski, R. Regional catchment classification with respect to low flow risk in a Pleistocene landscape. *J. Hydrol.* **2012**, *475*, 392–402. [CrossRef]

20. Liu, W.-C.; Yu, H.-L.; Chung, C.-E. Assessment of water quality in a subtropical Alpine Lake using multivariate statistical techniques and geostatistical mapping: A case study. *Int. J. Environ. Res. Public Health* **2011**, *8*, 1126–1140. [CrossRef] [PubMed]

21. Wang, Y.; Wang, P.; Bai, Y.; Tian, Z.; Li, J.; Shao, X.; Mustavich, L.F.; Li, B.-L. Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China. *J. Hydro-Environ. Res.* **2013**, *7*, 30–40. [CrossRef]

22. Berman, M.; Phatak, A.; Traylen, A. Some invariance properties of the minimum noise fraction transform. *Chemom. Intell. Lab. Syst.* **2012**, *117*, 189–199. [CrossRef]

23. Luo, G.; Chen, G.; Tian, L.; Qin, K.; Qian, S.-E. Minimum noise fraction versus principal component analysis as a preprocessing step for hyperspectral imagery denoising. *Can. J. Remote Sens.* **2016**, *42*, 106–116. [CrossRef]

24. Moore, F.; Rastmanesh, F.; Asadi, H.; Modabberi, S. Mapping mineralogical alteration using principal-component analysis and matched filter processing in the Takab area, north-west Iran, from ASTER data. *Int. J. Remote Sens.* **2008**, *29*, 2851–2867. [CrossRef]

25. Müller, B.; Bernhardt, M.; Schulz, K. Identification of catchment functional units by time series of thermal remote sensing images. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 5345–5359. [CrossRef]

26. Munyati, C. Use of principal component analysis (PCA) of remote sensing images in wetland change detection on the Kafue Flats, Zambia. *Geocarto Int.* **2004**, *19*, 11–22. [CrossRef]

27. Gupta, R.P.; Tiwari, R.K.; Saini, V.; Srivastava, N. A simplified approach for interpreting principal component images. *Adv. Remote Sens.* **2013**, *2*, 111–119. [CrossRef]

28. Parmentier, B. Characterization of land transitions patterns from multivariate time series using seasonal trend analysis and principal component analysis. *Remote Sens.* **2014**, *6*, 12639. [CrossRef]

29. Li, J.H.; Wang, S.S.; Zhou, F.Q. Time series analysis of long-term terrestrial water storage over Canada from GRACE Satellites using principal component analysis. *Can. J. Remote Sens.* **2016**, *42*, 161–170. [CrossRef]

30. De Almeida, T.I.R.; Penatti, N.C.; Ferreira, L.G.; Arantes, A.E.; do Amaral, C.H. Principal component analysis applied to a time series of MODIS images: The spatio-temporal variability of the Pantanal wetland, Brazil. *Wetl. Ecol. Manag.* **2015**, *23*, 737–748. [CrossRef]

31. Chen, G.; Qian, S.-E. Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis. *Can. J. Remote Sens.* **2008**, *34*, 447–454. [CrossRef]

32. Zabalza, J.; Ren, J.; Yang, M.; Zhang, Y.; Wang, J.; Marshall, S.; Han, J. Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 112–122. [CrossRef]

33. Liu, X.; Gao, L.; Zhang, B.; Zhang, X.; Luo, W. An improved MNF transform algorithm on hyperspectral images with complex mixing ground objects. In Proceedings of the 2008 Congress on Image and Signal Processing, Sanya, China, 27–30 May 2008; IEEE Computer Society: Washington, DC, USA, 2008; Volume 3, pp. 479–483.

34. Duveiller, G.; Lopez-Lozano, R.; Cescatti, A. Exploiting the multi-angularity of the MODIS temporal signal to identify spatially homogeneous vegetation cover: A demonstration for agricultural monitoring applications. *Remote Sens. Environ.* **2015**, *166*, 61–77. [CrossRef]

35. Xavier, L.; Becker, M.; Cazenave, A.; Longuevergne, L.; Llovel, W.; Filho, O.C.R. Interannual variability in water storage over 2003–2008 in the Amazon Basin from GRACE space gravimetry, in situ river level and precipitation data. *Remote Sens. Environ.* **2010**, *114*, 1629–1637. [CrossRef]

36. Food and Agriculture Organization of the United Nations-International Soil Reference and Information Centre (FAO-ISRIC). Soil and Terrain Database for Southern Africa (1:2 million scale). FAO Land and Water Digital Media Series 25, ISRC and FAO, Rome. Available online: http://geonode.isric.org/layers/geonode: soter_saf_map_unit#more (accessed on 23 November 2014).

37. Food and Agriculture Organization of the United Nations (FAO). The Multipurpose Africover Database for the Environmental Resources. Available online: http://www.fao.org/geonetwork/srv/en/main.search? title=africover%20landcover (accessed on 16 November 2014).

38. Burgess, N.D.; Butynski, T.M.; Cordeiro, N.J.; Doggart, N.H.; Fjeldså, J.; Howell, K.M.; Kilahama, F.B.; Loader, S.P.; Lovett, J.C.; Mbilinyi, B.; et al. The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. *Biol. Conserv.* **2007**, *134*, 209–231. [CrossRef]

39. Burgess, N.D.; Hales, J.D.A.; Ricketts, T.H.; Dinerstein, E. Factoring species, non-species values and threats into biodiversity prioritisation across the ecoregions of Africa and its islands. *Biol. Conserv.* **2006**, *127*, 383–401. [CrossRef]

40. Fisher, B.; Turner, R.K.; Burgess, N.D.; Swetnam, R.D.; Green, J.; Green, R.E.; Kajembe, G.; Kulindwa, K.; Lewis, S.L.; Marchant, R.; et al. Measuring, modeling and mapping ecosystem services in the Eastern Arc Mountains of Tanzania. *Prog. Phys. Geogr.* **2011**, *35*, 595–611. [CrossRef]

41. Schaafsma, M.; Morse-Jones, S.; Posen, P.; Swetnam, R.D.; Balmford, A.; Bateman, I.J.; Burgess, N.D.; Chamshama, S.A.O.; Fisher, B.; Green, R.E.; et al. Towards transferable functions for extraction of Non-timber Forest Products: A case study on charcoal production in Tanzania. *Ecol. Econ.* **2012**, *80*, 48–62. [CrossRef]

42. Madoffe, S.; Hertel, G.D.; Rodgers, P.; O'Connell, B.; Killenga, R. Monitoring the health of selected eastern arc forests in Tanzania. *Afr. J. Ecol.* **2006**, *44*, 171–177. [CrossRef]

43. Doggart, N.D.; Burgess, N.D. State of the Arc in 2005. *East. Arc Mt. Monit. Baselines* **2005**, *19*, 1–32.

44. Madulu, N.F. Environment, poverty and health linkages in the Wami River basin: A search for sustainable water resource management. *Phys. Chem. Earth* **2005**, *30*, 950–960. [CrossRef]

45. Wambura, F.J. Stream flow response to skilled and non-linear bias corrected GCM precipitation change in the Wami River Sub-basin. *Br. J. Environ. Clim. Chang.* **2014**, *4*, 389–408. [CrossRef]

46. Wambura, F.J.; Ndomba, P.M.; Kongo, V.; Tumbo, S.D. Uncertainty of runoff projections under changing climate in Wami River sub-basin. *J. Hydrol. Reg. Stud.* **2015**, *4*, 333–348. [CrossRef]

47. Jarvis, A.; Reuter, H.I.; Nelson, A.; Guevara, E. Hole-Filled SRTM for the Globe Version 4. The CGIAR-CSI SRTM 90m Database. 2008. Available online: http://srtm.csi.cgiar.org (accessed on 13 November 2014).

48. Weedon, G.P.; Balsamo, G.; Bellouin, N.; Gomes, S.; Best, M.J.; Viterbo, P. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.* **2014**, *50*, 7505–7514. [CrossRef]

49. Japan International Cooperation Agency (JICA). *The Study on Water Resources Management and Development in Wami/Ruvu Basin in the United Republic of Tanzania*; Japan International Cooperation Agency (JICA) on behalf of Tanzanian Ministry of Water: Dar es Salaam, Tanzania, 2013.

50. Wang, L.; Wang, Y.; Xin, J.; Li, Z.; Wang, X. Assessment and comparison of three years of Terra and Aqua MODIS Aerosol Optical Depth Retrieval (C005) in Chinese terrestrial regions. *Atmos. Res.* **2010**, *97*, 229–240. [CrossRef]

51. Mu, Q.; Zhao, M.; Running, S.W. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* **2011**, *115*, 1781–1800. [CrossRef]

52. Tasumi, M.; Hirakawa, K.; Hasegawa, N.; Nishiwaki, A.; Kimura, R. Application of MODIS land products to assessment of land degradation of Alpine Rangeland in Northern India with limited ground-based information. *Remote Sens.* **2014**, *6*, 9260–9276. [CrossRef]

53. Byrne, G.F.; Crapper, P.F.; Mayo, K.K. Monitoring land-cover change by principal component analysis of multitemporal landsat data. *Remote Sens. Environ.* **1980**, *10*, 175–184. [CrossRef]

54. Lee, J.; Verleysen, M. *Nonlinear Dimensionality Reduction Information Science and Statistics*; Springer: New York, NY, USA, 2007.

55. An, Q.; Wu, Y.Q.; Wang, J.H.; Li, Z.E. Assessment of dissolved heavy metal in the Yangtze River estuary and its adjacent sea, China. *Environ. Monit. Assess.* **2010**, *164*, 173–187. [CrossRef] [PubMed]

56. Huang, J.; Ho, M.; Du, P. Assessment of temporal and spatial variation of coastal water quality and source identification along Macau peninsula. *Stoch. Environ. Res. Risk Assess.* **2011**, *25*, 353–361. [CrossRef]

57. Lu, A.; Wang, J.; Qin, X.; Wang, K.; Han, P.; Zhang, S. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. *Sci. Total Environ.* **2012**, *425*, 66–74. [CrossRef] [PubMed]

58. Chen, Y.; McPhedran, M.N.; Perez-Estrada, L.; El-Din, M.G. An omic approach for the identification of oil sands process-affected water compounds using multivariate statistical analysis of ultrahigh resolution mass spectrometry datasets. *Sci. Total Environ.* **2015**, *511*, 230–237. [CrossRef] [PubMed]

59. Lischeid, G.; Steidl, J.; Merz, C. Functional versus trend analysis to assess anthropogenic impacts on groundwater heads. *Grundwasser* **2012**, *17*, 79–89. [CrossRef]

60. Johansson, J. Introduction to Scientific Computing in Python. 2014. Available online: http://www-star.st-and.ac.uk/~pw31/CompAstro/IntroToPython.pdf (accessed on 6 January 2015).

61. Oliphant, T.E. *A Guide to NumPy*; Trelgol Publishing: Spanish Fork, UT, USA, 2006; Volume 1.

62. Oliphant, T.E. Python for scientific computing. *Comput. Sci. Eng.* **2007**, *9*, 10–20. [CrossRef]

63. Mann, H.B. Nonparametric tests against trend. *Econometrica* **1945**, *13*, 245–259. [CrossRef]

64. Kendall, M.G. *Rank Correlation Methods*; Griffin: London, UK, 1975.

65. Sen, P.K. Estimates of the regression coefficient based on Kendall's Tau. *J. Am. Stat. Assoc.* **1968**, *63*, 1379–1389. [CrossRef]

66. Yue, S.; Pilon, P.; Phinney, B.; Cavadias, G. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Processes* **2002**, *16*, 1807–1829. [CrossRef]

67. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [CrossRef]

68. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1945**, *1*, 80–83. [CrossRef]

69. Yang, Y.; Long, D.; Guan, H.; Liang, W.; Simmons, C.; Batelaan, O. Comparison of three dual-source remote sensing evapotranspiration models during the MUSOEXE-12 campaign: Revisit of model physics. *Water Resour. Res.* **2015**, *51*, 3145–3165. [CrossRef]

70. Lewandowski, J.; Lischeid, G.; Nuetzmann, G. Drivers of water level fluctuations and hydrological exchange between groundwater and surface water at the lowland River Spree (Germany): Field study and statistical analyses. *Hydrol. Process.* **2009**, *23*, 2117–2128. [CrossRef]

71. Skarbek, C. A review of endemic species in the Eastern Arc Afromontane Region: Importance, inferences, and conservation. *Macalester Rev. Biogeogr.* **2008**, *1*, 3.

72. Fjeldså, J.; Kiure, J.; Doggart, N.; Hansen, L.; Perkin, A. Distribution of highland forest birds across a potential dispersal barrier in the Eastern Arc Mountains of Tanzania. *Steenstrupia* **2010**, *32*, 1–43.

73. Kijazi, A.L.; Reason, C.J.C. Analysis of the 2006 floods over northern Tanzania. *Int. J. Climatol.* **2009**, *29*, 955–970. [CrossRef]

74. Mapande, A.T.; Reason, C.J.C. Interannual rainfall variability over Western Tanzania. *Int. J. Clim.* **2005**, *25*, 1355–1368. [CrossRef]

75. Forestry and Beekeeping Division (FBD). *Forest Area Baseline for the Eastern Arc Mountains*; Mbilinyi, B.P., Malimbwi, R.E., Shemwetta, D.T.K., Songorwa, E.Z., Katani, J.Z., Kashaigili, J., Eds.; Conservation and Management of the Eastern Arc Mountain Forests, Forestry and Beekeeping Division: Dar es Salaam, Tanzania, 2006.

76. Green, J.M.H.; Burgess, N.D.; Green, R.E.; Madoffe, S.S.; Munishi, P.K.T.; Nashanda, E.; Turner, R.K.; Balmford, A. Estimating management costs of protected areas: A novel approach from the Eastern Arc Mountains, Tanzania. *Biol. Conserv.* **2012**, *150*, 5–14. [CrossRef]

77. Bayliss, J.; Schaafsma, M.; Balmford, A.; Burgess, N.D.; Green, J.M.H.; Madoffe, S.S.; Okayasu, S.; Peh, K.S.H.; Platts, P.J.; Yu, D.W. The current and future value of nature-based tourism in the Eastern Arc Mountains of Tanzania. *Ecosyst. Serv.* **2014**, *8*, 75–83. [CrossRef]

78. Rovero, F.; Menegon, M.; Fjeldså, J.; Collett, L.; Doggart, N.; Leonard, C.; Norton, G.; Owen, N.; Perkin, A.; Spitale, D.; et al. Targeted vertebrate surveys enhance the faunal importance and improve explanatory models within the Eastern Arc Mountains of Kenya and Tanzania. *Divers. Distrib.* **2014**, *20*, 1438–1449. [CrossRef]

79. Schaafsma, M.; Morse-Jones, S.; Posen, P.; Swetnam, R.D.; Balmford, A.; Bateman, I.J.; Burgess, N.D.; Chamshama, S.A.O.; Fisher, B.; Freeman, T.; et al. The importance of local forest benefits: Economic valuation of Non-Timber Forest Products in the Eastern Arc Mountains in Tanzania. *Glob. Environ. Chang.* **2014**, *24*, 295–305. [CrossRef]

80. Swetnam, R.D.; Fisher, B.; Mbilinyi, B.P.; Munishi, P.K.; Willcock, S.; Ricketts, T.; Mwakalila, S.; Balmford, A.; Burgess, N.D.; Marshall, A.R.; et al. Mapping socio-economic scenarios of land cover change: A GIS method to enable ecosystem service modelling. *J. Environ. Manag.* **2011**, *92*, 563–574. [CrossRef] [PubMed]

81. Tabor, K.; Burgess, N.D.; Mbilinyi, B.P.; Kashaigili, J.J.; Steininger, M.K. Forest and woodland cover and change in coastal Tanzania and Kenya, 1990 to 2000. *J. East Afr. Nat. Hist.* **2010**, *99*, 19–45. [CrossRef]