

Article

## Centered Log-Ratio (clr) Transformation and Robust Principal Component Analysis of Long-Term NDVI Data Reveal Vegetation Activity Linked to Climate Processes

Faith K. Muriithi

Department of Earth and Environmental Studies, Montclair State University, 1 Normal Avenue, Montclair, NJ 07043, USA; E-Mail: justusf2@montclair.edu; Tel.: +1-973-674-0480

Academic Editors: Aondover Tarhule and Zewdu T. Segele

Received: 14 November 2014 / Accepted: 4 January 2015 / Published: 13 January 2015

---

**Abstract:** Predicting the future climate and its impacts on the global environment is model based, presenting a level of uncertainty. Alternative robust approaches of analyzing high volume climate data to reveal underlying regional and local trends are increasingly incorporating satellite data. This study uses a centered log-ratio (clr) transformation approach and robust principal component analysis (PCA), on a long-term Normalized Difference Vegetation Index (NDVI) dataset to test its applicability in analyzing large multi-temporal data, and potential to recognize important trends and patterns in regional climate. Twenty five years of NDVI data derived by Global Inventory Modeling and Mapping Studies (GIMMS) from 1982 to 2006 were extracted for 88 subwatersheds in central Kenya and statistically analyzed. Untransformed (raw) and clr transformed NDVI data were evaluated using robust PCA. The robust PCA compositional biplots of the clr transformed long-term NDVI data demonstrated the finest spatial-temporal display of observations identifying climate related events that impacted vegetation activity and observed variations in greenness. The responses were interpreted as normal conditions, El Niño Southern Oscillation (ENSO) events of El Niño and La Niña, and drought events known to influence the moisture level and precipitation patterns (high, low, normal) and therefore the level of vegetation greenness (NDVI value). More drought events (4) were observed between 1990 and 2006, a finding corroborated by several authors and linked to increasing climate variability. Results are remarkable, emphasizing the need for appropriate data transformation prior to PCA, dealing with huge complex datasets, to enhance pattern recognition and meaningful interpretation of results. Through improved analysis of past data, uncertainty is decreased in modeling future trends.

**Keywords:** PCA; data transformation; vegetation; climate; drought; ENSO

---

## 1. Introduction

Global, regional and local scale studies of vegetation dynamics widely utilize synoptic satellite data since it provides the fine spatial and temporal details necessary to understand plant phenological activities. Even so, obtaining, processing and analyzing remote sensing data can be a challenge for non-remote sensing specialists and amateurs due to complexity of computer based algorithms and mathematical treatment involved [1]. In the early 1990s, simpler techniques for exploring data in a spatial context were often ignored in favor of elegant formal modeling [2], and even today, this perception still persists among the pure model developers. Complex computer based algorithms can limit data exploration and application by the wider group of interested researchers. The long-term data provided by long mission satellites e.g., the Landsat series, National Oceanic and Atmospheric Administration (NOAA)—Advanced Very High Resolution Radiometer (AVHRR) series has wide-ranging applications across disciplines, and primarily modeling of earth systems processes and their interaction.

To meet the challenge, research teams such as the group responsible for processing and provision of NOAA Global Inventory Modeling and Mapping Studies (GIMMS) 1982–2006 [3] provide global long-term data in different formats e.g., tagged image files (.tif) that are easier to use and compatible with a wide range of geographical information systems and spatial data processing software (ArcGIS, Erdas Imagine, GRASS GIS, *etc.*). This software provide user friendly interface, reducing the level of difficulty in processing digital image data, thereby supporting a broader scientific research interest.

Despite this advancement, the choice of statistical methodology, and the intuitive interpretation of the results require skill and expertise that is often accumulated over years. Multivariate statistical analysis such as principal component analysis (PCA) has a broad application in analysis of complex datasets derived from remote sensing images. PCA often involve transformation of multispectral bands such that the new bands (also called components) are uncorrelated [4], and capture the essential information from the huge dataset. These are ordered in terms of the amount of image variation they explain [5]. Furthermore, the high volume and dimensionality of remote sensing data particularly the hyperspectral images present a challenge to researchers [6], and as a result, computational methods that reduce the processing burden, such as PCA, become very useful [7].

Of particular interest to this work is the Normalized Difference Vegetation index (NDVI) data, commonly used as an indicator of vegetation greenness, health and vigor [8–10]. The sun's electromagnetic energy arriving on plant surface is absorbed/transmitted or reflected. Plants utilize the visible energy in photosynthesis; however, radiation emanating from the near infrared energy is mainly reflected by the internal cell arrangement of the leaves. These two types of energy have provided a breakthrough in terrestrial research focused on vegetation dynamics and health including drought, stress indication, carbon output, climate modeling, and disease response [11]. NDVI highly correlates with such parameters as green leaf biomass, the density of chlorophyll contained in plants [9] and leaf area index (LAI), all which are parameters that gauge vegetation activity and health [12]. Though NDVI has limitations and challenges, e.g., in presence of scattered vegetation canopy, great amount of background

spectral response can be picked by the radiometers, leading to spurious ratio; presence of aerosols and particulate matter in the atmosphere (requiring rigorous data cleaning before use), saturating way before maximum leaf or plant biomass is obtained (providing nonlinear relationship with leaf area, plant cover), these challenges still remain relatively negligible as compared to its usefulness. NDVI is presented as the ratio of the difference between near-infrared reflectance and red visible reflectance to their sum. The vegetated surfaces take relative values between 0.1 and 0.7 [13] while non-vegetated surfaces have values close to zero. Such ratio values constrained to unity are described by [14,15] as closed values or compositional data values, and discussed at length [16,17]. Other examples of such datasets include geochemical data expressed in part per million (ppm), milligram per liter (mg/L) or proportions expressed as percentages (%).

To effectively undertake PCA of such closed data, log-ratio transformation is highly recommended which “opens” or un-constrains the data [16,18] to reveal underlying/inherent patterns in the data structure in a more vivid manner providing straight forward interpretation. Following principal component analysis (robust), the resulting compositional biplots are considered a standard tool to present multivariate data in a planar view, providing a deeper investigation of the data structure. The arrangement and display of rays (loadings) is based on the overall correlation matrix [19]. In the study, compositional biplots were used to explore the long-term vegetation condition and the spatio-temporal relationship with the regional climate variability in central Kenya. Compared to bivariate scatterplots, PCA biplots provide an overall planar view of the multivariate relationships of observations.

In this study, 88 sub-watersheds in central Kenya are used to demonstrate effectiveness and the importance of centered log-ratio transformation (clr) over untransformed (raw) data in principal component analysis (PCA), a broadly accepted multivariate approach. Changes in climate impact long term rainfall patterns, and seasonal variations [20–25], which directly impact vegetation growth and vigor. The arid and semi-arid lands (ASALs) in Africa are considered most vulnerable to climate variability due to its impact on primary productivity and fiber production that supports wildlife, pastoral livestock, agricultural yield, *etc.* Studies indicate that a decline in water and pasture resources in the arid and semi-arid regions of Kenya relates to recurrent and prolonged droughts [26], and also influences vegetation growth. Lines of evidence indicate that increased climate variability will likely impact the length of crop growing period over East Africa causing decline by 5%–20% by 2020 and over 20% by 2050 [27,20]. This will not only impact agricultural production [28] but also induce changes in vegetation dynamics and reduced total plant biomass, causing loss of fiber for livestock and wildlife. Methods that augment data analysis and by a wider group of researchers, while providing meaningful results regionally and locally are desirable in designing acceptable mitigation and adaption mechanisms at the community level.

Sub-watersheds within central Kenya are highly productive agrizones, which attract increased population growth and land fragmentation into small parcels per household (~2 ha). This exerts pressure on the natural resource base creating environmental changes with profound implication owing to increased land-use and land-cover transformations to suit the needs of the expanding communities *i.e.*, residential, commercial. A recent study [29] indicates that population dynamics play an important role in examining vulnerability to climate change in the region and therefore demographic data needs to be included in designing mitigative and adaptive approaches to counter variability in climate change. Intensive large-scale commercial horticulture practices in central Kenya [30,31], presents a growing

challenge to sustainable resource use and management. While agriculture intensification is widely viewed as a better option to enhance food security and reduce malnutrition in sub-Saharan nations [32,33], research on the impacts of intensive activities on watershed resources is still developing, and reveals negative environmental impacts. The mono-crop systems used by large-scale commercial horticulture farms may place such systems at high vulnerability to climate change; however, unlike small-scale farmers who have limited technology, little access to extension services, and market disadvantages [34]; large-scale farms may have resources to cope with increased changes. The described issues identify the region as a perfect case study.

The study demonstrates the necessity of (i) closed data transformation in pattern recognition and interpretation of long-term Normalized Difference Vegetation Index (NDVI) dataset (ii) compositional PCA biplots in explaining underlying variability in regional vegetation greenness, comparing the clr transformed and untransformed data.

## 2. Methods and Data

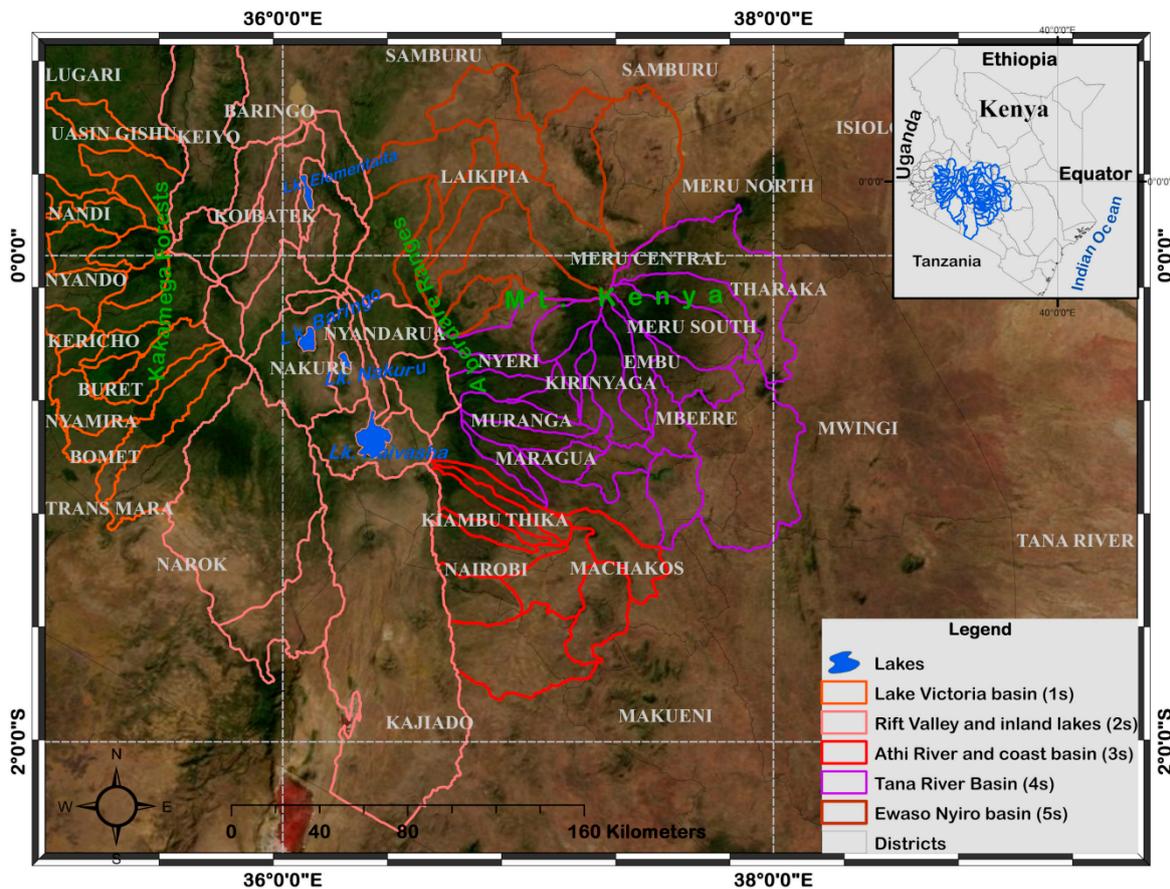
### 2.1. Study Area

The study area covered 88 sub-watersheds in the central highland of Kenya, confined within: 34°55'51.88"E, 0°55'10.51"N (upper-left); 38°07'09.91"E, 2°23'03.73"S (lower-right) and an area of 81,607.26 km<sup>2</sup> (Figure 1). These were chosen because of the pronounced variability in climate, vegetation, and rainfall and therefore demonstrate the spatio-temporal variations in phenological response to changing climate and environmental variables. The area has six diverse agro-ecological zones with varying physical environment of an almost equatorial type of climate in upper-lands and semi-arid and arid environments in the lowlands [35]. The upper-lands occur in the upper northwest region of the study area and are generally humid to sub humid, often presenting thick forest vegetation (Kakamega forest—See Figure 1). The lowlands show strong gradients in terms of average rainfall, temperature, and vegetation characteristics. There is high variability in rainfall within the study area, which occurs in two seasons, April–May (long rains) and October–November (short rains). Around the Aberdare ranges (Figure 1) the average annual rainfall is about 1350 mm/year. However, this reduces substantially to ~600 mm/year near Lake Naivasha. Another interesting feature to the east of the area is Mt. Kenya which stands at 5199 m above sea level. The mountain causes a strong shadow effect that highly influences climate conditions of the adjoining plateau land to more of a semi-arid to arid region, with rainfall of ~450 to 750 mm/year [36]. Due to low average annual rainfall, vegetation to the eastern and southern (760 mm/year) regions is dominated by woody vegetation of *Acacia* and *Themeda* grass cover that mainly occur in savanna environments.

#### 2.1.1. Satellite Data

Global Inventory Modeling and Mapping Studies (GIMMS) NDVI data extracted from AVHRR (1982–2006) for 88 sub-watershed in central Kenya was used. Long-term time series data are essential in detection of vegetation trends [37]. The data provides biweekly maximum value composited NDVI data with an 8-km spatial resolution from 1982 to 2006 [3] and was recently extended to 2010. Healthy green vegetation has the high NDVI values (>0.7 for forests; 0.3–0.4 for grasslands) while surfaces

without vegetation, such as bare soil, water, snow, ice or clouds typically have low NDVI values near zero. The long-term standardized data are derived from NOAA AVHRR imagery using the NOAA satellite series 7, 9, 11, 12, and 16. This version of the GIMMS NDVI dataset is corrected through a series of processing steps to alleviate known limitations of the AVHRR measurements induced by inter-sensor calibration differences, orbital drift, cloud cover, solar zenith and viewing angles differences, volcanic eruptions, and other atmospheric contaminations [38]. The excellent spatial coverage and relatively long-term observations by this NDVI dataset enable trend analyses. Work by [37] shows that the GIMMS data is most accurate AVHRR-NDVI dataset for assessing vegetation variability and trends. Earlier versions of the NDVI dataset have been widely used in detecting vegetation growth change [39].



**Figure 1.** Study area showing the geographical location of the sub-watersheds (within the five main basins in Kenya) and important physical features in central Kenya. Other features include; the equator that crosses the country at 0°0'00", and the Indian Ocean coastline.

2.1.2. Sub-Watershed Data

The sub-watershed data layer used in the study was created by the World Research Institute (WRI) from the 1992 Kenya National Water Master Plan, and is publically available [40]. Each sub watershed is assigned a unique code composed of a number and two letters, e.g., 5AB. The number represents one of the five river basins in Kenya, *i.e.*, Lake Victoria (1); Rift valley and inland lakes (2); Athi River and coast (3); Tana River (4) and Ewaso Ng’iro (5) in which the sub-watershed resides. The first letter (in

this case A) represents the sub catchment, while the second letter (in this case B) represents the major river branch in the particular sub-catchment. Given that a sub catchment may have several major river branches, the second letters may continue gradually to as many as there are major river branches in the sub-catchment, for instance, 5AA, 5AB, 5AC, 5AD, 4BA, 4BB, *etc.*

## 2.2. Data Analysis

### 2.2.1. Data Preparation—NDVI Data Extraction

Maximum value composite NDVI (which is a maximum daily NDVI value during the 15-day period) minimizes atmospheric effects and cloud contamination effects. Monthly biweekly images were processed for the period of study (25 years) and for all the 88 sub-watersheds. The maximum NDVI values of the first 15 days of each month were utilized in generating monthly data. The monthly values were then aggregated to determine the annual mean value. To reduce the impact of bare and sparsely vegetated pixels on the NDVI trend, all pixels in sub-watershed (masks) were used in zonal statistics to derive a single maximum NDVI for the first 15 days. The above steps were carried out using the zonal statistics tool in ArcGIS 10.1. In total, 25 years of bi-monthly maximum value composite data for 88 sub-watersheds were processed resulting to 600 image subsets for the 88 study units.

### 2.2.2. Centered Log-Ratio (clr) Transformation and Robust Principal Component Analysis (PCA)

Essentially, composition data are measures of proportions, percentages, parts per million, *etc.*, which sums up to a constant value—also referenced as closed data [14,15]. Such data is constrained [19] and require a transformation to remove the constraint, thereby making it usable in context of standard statistical approaches of principal component analysis (PCA). A transformation-based methodology to deal with such data was proposed in the early 1980s [14], largely due to the recognition that compositions provided information only about the relative magnitudes of their components, and hence a need for data analysis focused on the ratios between components [41]. Two main transformations are commonly used today; the additive log-ratio transformation and the centered log-ratio. Given that the NDVI data is presented as the ratio of the difference between near-infrared (NIR) reflectance and red (R) visible reflectance to their sum [42], it was considered closed data suitable for clr transformation. The magnitude of NDVI values *i.e.*, 0 to 1, provide an indication of vegetation condition (low to high, implying non vegetative grounds to dense green vegetation).

The current study utilized the centered log-ratio transformation, mathematically expressed as:

$$\text{clr}(x) = (\log(x_1/g(x)), \dots, \log(x_D/g(x)))$$

where  $x$  represents the composition vector,  $g(x)$  is the geometric mean of the composition  $x$ , and  $x_D$  is Euclidean distances between the individual NDVI variables. The clr transformation was carried out using CoDaPack, a freely available and easy to use, Excel based software for compositional data transformation [43]. However, there are other numerous, more specialized packages utilized in compositional data analysis. For instance; robCompositions, which compute robust estimation for compositional data [44], the isometric logratio (ilr) [18], and Compos Analysis, which is an Add-On of

the centered MANOVA analysis [45,46]. Following the centered log-ratio transformation, the results were presented as a clr plot, a graphic showing NDVI variables for the sub-watersheds in three dimensions.

The use of centered log-ratio transformation approach on the long term NDVI dataset, besides the regular PCA on untransformed NDVI values, is unique. It tests the algorithms applicability in analyzing large multi-temporal data datasets, and if it can recognize trends and patterns in regional climate events.

PCA (robust) of un-transformed (raw) data was carried out on the derived NDVI data. Principal component analysis (PCA) using the robust option was carried out in JMP software, generating the principal components (PCs) and compositional biplots (load and scores).

### 3. Results and Discussion

The results demonstrate a pronounced contrast in the arrangement and display of variables between the untransformed and the centered logratio (clr) transformed NDVI data as shown in Figures 2 and 3. The compositional biplots, *i.e.*, the loading plots and score plots in Figure 2B and 3B, of the clr transformed data, express reasonable and substantial information on interannual variability in vegetation condition in the sub-watersheds. The variability in annual NDVI in the subwatersheds for the last 25 years is explained by 2 principal components (PCs) in both scenarios (56.3% and 50% respectively). The untransformed data exhibits a slightly higher variability (Figure 2A~56.3%) but very little temporal information is obtained regarding similarities or differences in vegetation condition in sub-watersheds. The untransformed data exhibits a behavior commonly referred to as closure [14,19], and is shown to often influence compositional data regardless of the number of elements used and/or their magnitude, and therefore the need for data transformation before a PCA analysis [16].

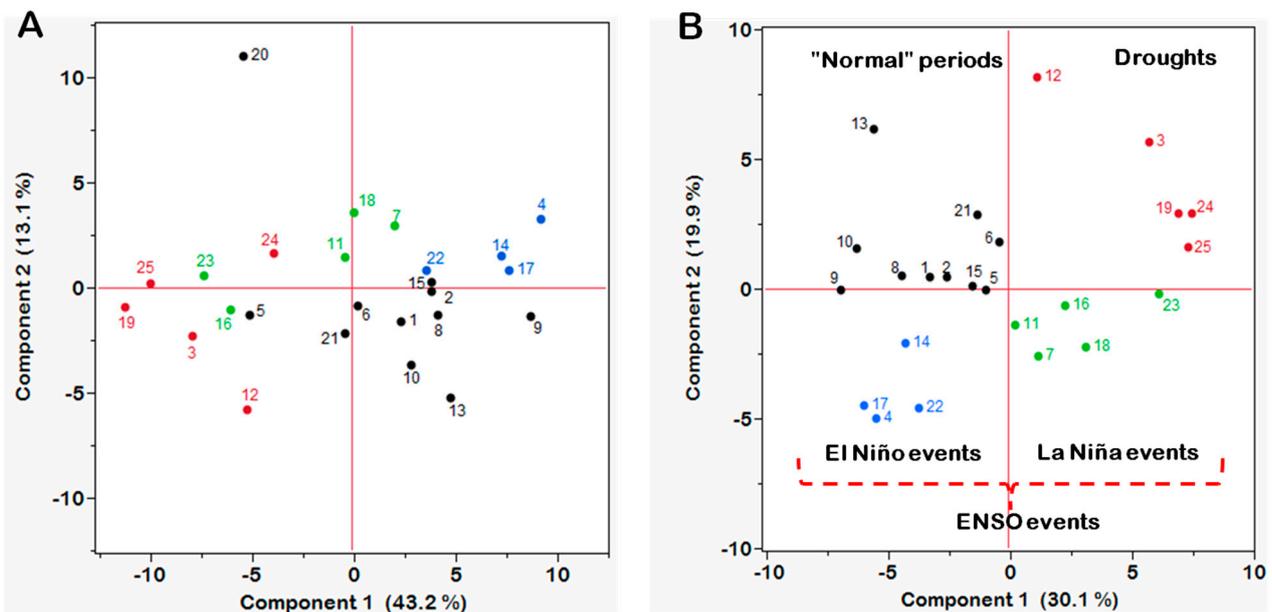
From the PCA (robust) results of the clr transformed data (Figures 2B and 3B), the longterm observations of inter-annual NDVI (1982–2006) are reduced to two PCs, that explain 50% of variability in vegetation greenness. The observations are evidently separated into four main events, as that is easier to visualize. Moreover, years with similar vegetation responses tend to deviate from normal conditions and appear as tighter clusters in separate quadrants (composition space). These events primarily represent climate linked processes of ENSO (El Niño and La Niña) that are known to impact the East Africa rainfall and vegetation dynamics.

Generally, the results resonate with a broad literature indicating the connection between vegetation greenness and rainfall patterns in East Africa with the coupled ocean—Atmosphere phenomena of El Niño Southern Oscillation (ENSO) and the Indian Ocean Dipole [47,48]. ENSO forcing is widely linked to atypical rainfall patterns [49] causing higher than normal photosynthetic activity across East and southern Africa [50]. Until recently the role of the Indian Ocean Dipole in driving photosynthetic anomalies was generally uncertain. Emerging studies [47,51] indicate a strong coupled forcing of both climate-Atmosphere-Ocean processes that are highly connected to unusual rainfall and anomalies in vegetation greenness in East Africa.

#### 3.1. Drought Events Impacting Vegetation Condition

The years 1984 (3), 1993 (12), 2000 (19), 2005 (24) and 2006 (25) were clustered in the upper right quadrant (Figure 2B). These were identified as drought events that occurred in the region during the 25 years of study and were confirmed by a broad array of literature [49,50,52–55] among many other

studies. Drought limits the amount of soil moisture available for plants, limiting net photosynthesis [56]. Drought impact on vegetation is highly dependent on soil type, vegetation type, temperature, intensity, duration, and the overall health and vigor of the vegetation cover prior to the drought, *etc.* Though plants store food reserves in the event of droughts, in extreme drought condition, the stored food reserves may not suffice and this increases plant’s susceptibility to damage during prolonged droughts. Moreover, reduced aboveground photosynthetic biomass (green leaves, green stems, and shoots) limits chlorophyll production causing decline in vegetation and overall greenness and therefore a low NDVI. Plants with healthy root networks and sufficient carbohydrate reserves fare much better in the event of, and even after drought compared to the resource limited plants [57].

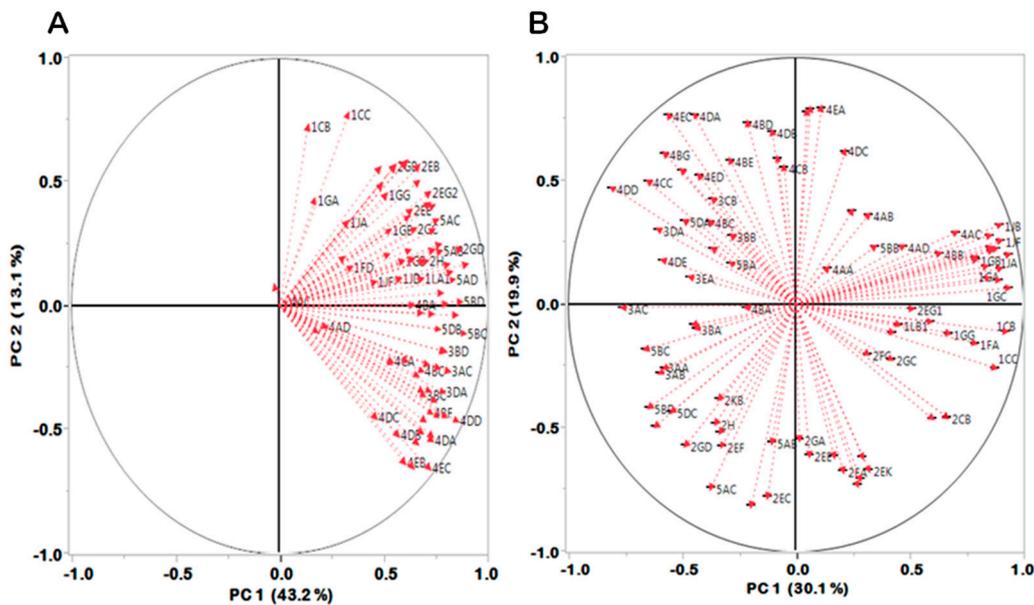


**Figure 2.** Score plots of (A) untransformed (raw) NDVI data, (B) centered log-ratio (clr) transformed NDVI data. The numbers represent years chronologically starting from 1982 (1) to 2006 (25). Similar color scheme represent like events/phenomenon that occurred in the particular year.

These provide plausible explanation to the observed spatial arrangement of rays in Figure 3B, Figure 4. Distinct spatial clustering of longer rays (high variability) was observed in the upper right quadrant with sub-watersheds (1s) which are essentially thick canopy forest of the Kakamega forest, a few sub-watersheds surrounding Mt. Kenya forests, and the Aberdare ranges in central Kenya (see Figure 1 of study area). Generally, the spatial variability of vegetation activity across the study area is not homogeneous, a characteristic that could be linked to differences in ecosystems and the types of vegetation they support e.g., grasslands, wooded savanna, agriculture among other uses and therefore the distinct arrangement of rays (Figures 3B and 4) in the compositional space.

Studies show that the 2005–2006 drought, like other previous droughts, severely impacted pastoral communities and farming, while the rate of malnutrition increased sharply particularly in the north-east Kenya [54]. Lack of grazing pastures leads to the death of thousands of livestock among the pastoral communities [58]. Droughts are becoming more frequent, allowing less time for recovery in between

droughts, and increasing the vulnerability of local populations. Based on this study, four droughts occurred in the period between 1990 and 2006.



**Figure 3.** Loading plots of (A) untransformed data—Raw; (B) centered log-ratio (clr) transformed annual NDVI data. In the untransformed data, the rays representing the 88 sub-watersheds are clustered towards one direction. After transformation, substantial spread of rays in all directions was attained, and with smaller clusters of sub-watersheds showing varied response to changes in vegetation greenness. Longer and closely packed rays in upper right (UR) quadrant indicate a remarkable variation and correlation in NDVI by sub-watersheds within the Lake Victoria Basin (1s).

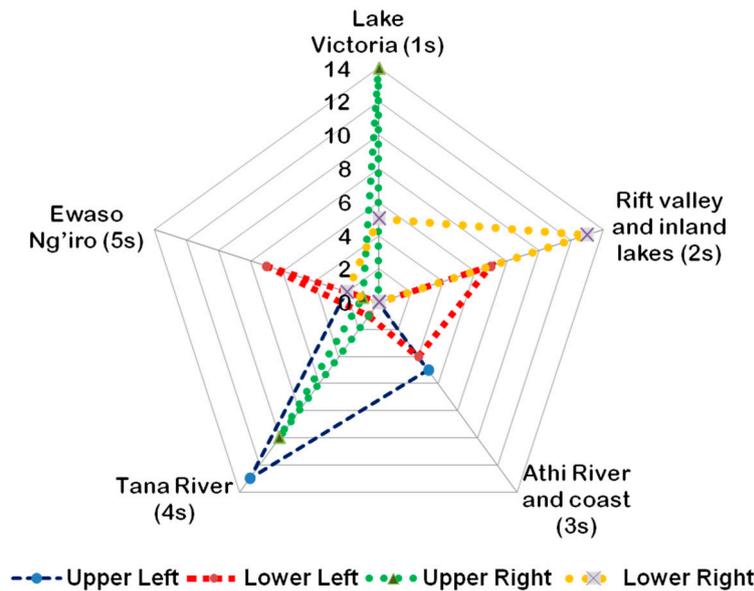
### 3.2. El Niño Southern Oscillation (ENSO)

The El Niño Southern Oscillation (ENSO)—which includes both El Niño and La Niña events—is the coupled large-scale ocean-atmosphere process involving changes in sea-surface temperature (SSTs) across the central and eastern tropical Pacific. However, for the East African region, recent studies show that on timescales beyond the decadal, the Indian Ocean drives East African rainfall variability by altering the local Walker circulation, whereas the influence of the Pacific Ocean is minimal [47]. These processes impact weather patterns, causing changes in rainfall frequency and intensity, which impact vegetation growth characteristics [59,60].

#### 3.2.1. El Niño—Extremely High Rainfall Events

The lower left quadrant (Figure 2B) illustrate a cluster of four years *i.e.*, 1985(4), 1995(14), 1998(17) and 2003(22), which represent years that had unusual heavy rainfall caused by El Niño. Generally, heavy rainfall promote higher than normal above ground primary productivity [61] leading to high NDVI values. Studies indicate that wet conditions in the coastal East Africa are closely linked to cool sea surface temperatures (SSTs) in the eastern Indian Ocean and warm sea surface temperatures (SSTs) in

the western Indian Ocean, atmospheric processes that cause ascending atmospheric circulation over East Africa and enhanced rainfall [47].



**Figure 4.** A radar plot illustration detailing the distribution of the sub-watersheds within the main five basins, and within the loading plot quadrants (Figure 3B). Sub-watersheds within the Lake Victoria Basin (1s) were common in the upper (UR) quadrant. Sub-watersheds in lower right (LR) quadrant were predominantly from Rift valley and inland lakes basin (2s); sub-watersheds in upper left (UL) quadrants were primarily from Tana River basin (4s) and few from Athi River and coast basin (3s); sub-watersheds in lower left (LL) quadrant were mainly from Ewaso Ng’iro (5s), and few others from Rift valley and inland lakes basin (2s).

### 3.2.2. La Niña Events—Cool Processes

In Figure 2B, the lower right quadrant indicate an obvious separation of the years 1988 (7), 1992 (11), 1997 (16), 1999 (18) and 2004 (23), interpreted as atypical events coinciding with La Niña events in East Africa region. Recent research shows that the Pacific Ocean has minimal influence on the ENSO pattern in East Africa, and that during La Niña events, the cold sea surface temperatures in the western Indian Ocean and warmer in the East Indian Ocean is what causes drought in the East Africa region. La Niña events cause poor short rains in East Africa, which are typically in the months of October–December [62,63]. Reduced moisture level limit plants physiological processes, leading to reduced leaf biomass and overall plant vigor. From the study, it appears that this type of response comes right before the drought period. It is plausible to think that during La Niña years, plants utilize most of stored reserves to promote growth; however, this is short-lived when rains fail leading to decline in vegetation and therefore low NDVI.

Other explanation supporting the observed separation of La Niña years from El Niño and drought events based on vegetation index is elaborated in [50]. They indicate that transition from El Niño to La Niña event influence vegetation dynamic in East Africa during the 1997–1998 El Niño events. Above normal greenness and patterns of above normal rainfall were observed over East Africa, which reversed

the decline in rainfall patterns and decline in vegetation greenness during La Niña. While their study used satellite images from 1998 to 2006, this study examines a longer record of data, defined regionally by sub-watersheds, and the findings are in agreement.

Besides vegetation characteristics, ENSO cycles have also important consequences for agriculture (whose yield thrives in high rainfall), livestock farming (enough pasture, fiber and thus increased reproduction), flooding, and also disease prevalence (occurrence of Malaria, Rift Valley Fever—1997/1998) in the arid and semi-arid areas of the country.

### 3.3. Normal/Neutral Conditions

Results in the upper left quadrant represent years where normal climate conditions prevailed, with no unusual events that may have impacted normal vegetation growth. They are generally referred to as “normal” years, where vegetation undergoes normal phenological processes/activity. The study region supports different plant ecosystems, among them the savanna ecosystems which responds rapidly to changing precipitation patterns and moisture levels.

## 4. Conclusions

The study demonstrates that robust principal component analysis of centered log-ratio transformed NDVI data provides meaningful and interpretable findings related to long-term climate processes influencing vegetation greenness in central Kenya unlike the un-transformed data.

Spatial and temporal isolation of tighter clusters of observations illustrated by the compositional biplots of clr transformed data allowed a straight forward explanation. Based on the finding, four main climate-related processes *i.e.*, normal/neutral years, drought events, ENSO (El Niño and La Niña events) affect vegetation dynamics in the study area, a conclusion commonly agreed upon as influencing rainfall patterns and vegetation activity in east Africa. The almost perfect match of events to known climatic condition the region is remarkable, further emphasizing the need for suitable data transformation before PCA.

Between 1982 and 2006, five drought events occurred in the study area, an observation that many scientists believe is primarily connected to increasing long-term climate variability and influenced by anthropogenic activities *e.g.*, deforestation and land degradation, greenhouse gas emissions, *etc.* The findings are important primarily because the region is an important socio-economic engine in terms of agricultural productivity, pastoral herding, resource availability, *etc.*, which could be vulnerable to increased atypical events.

## Acknowledgments

The author would like to thank Montclair State University for the graduate scholarship, and the Geological Society of America (GSA) for the graduate student research grant award that provided partial support towards accomplishment of this work.

## Conflicts of Interest

The author declares no conflict of interest.

## References

1. Richards, J.A. *Remote Sensing Digital Image Analysis*, 5th ed.; Springer-Verlag: Berlin, Germany, 2013.
2. Goodchild, M.; Haining, R.; Wise, S. Integrating GIS and spatial data analysis: Problems and possibilities. *Int. J. Geogr. Info. Syst.* **1992**, *6*, 407–423.
3. Tucker, C.J.; Pinzon, J.E.; Brown, M.E.; Slayback, D.; Pak, E.W.; Mahoney, R.; Vermote, E.; El Saleous, N. An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data. *Int. J. Remote Sens.* **2006**, *26*, 4485–4498.
4. Lhermitte, S.; Verbesselt, J.; Verstraaten, W.W.; Coppin, P. A comparison of time series similarity measures for classification and change detection of ecosystems dynamics. *Int. Remote Sens. Environ.* **2011**, *115*, 3129–3152.
5. Eastman, R.J. *IDRISI Selva Guide to GIS and Remote Sensing*; Clarklabs, Clark University: Worcester, MA, USA, 2012; p. 322.
6. Omkar, S.N.; Sivaranjani, V.; Senthilnath, J.; Mukherjee, S. Classification of hyperspectral data. *Int. J. Aero. Innov.* **2010**, *2*, 157–163.
7. Hyvarinen, A.; Oja, E. Independent component analysis: Algorithms and applications. *Neural Netw.* **2000**, *13*, 411–430.
8. Forkel, M.; Carvalhais, N.; Verbesselt, J.; Mahecha, M.D.; Neigh, C.S.R.; Reichstein, M. Trend change detection in NDVI time series: Effects of inter-annual variability and methodology. *Remote Sens.* **2013**, *5*, 2113–2144.
9. Sellers, P.J. Canopy reflectance, photosynthesis and transpiration. *Int. J. Remote Sens.* **1985**, *6*, 1335–1372.
10. Mennis, J. Exploring relationships between ENSO and vegetation vigour in the south-east USA using AVHRR data. *Int. J. Remote Sens.* **2001**, *22*, 3077–3092.
11. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510.
12. Wang, T.; Kou, X.; Xiong, Y.; Mou, P.; Wu, J.; Ge, J. Temporal and spatial patterns of NDVI and their relationship to precipitation in the Loess Plateau of China. *Int. J. Remote Sens.* **2010**, *31*, 1943–1958.
13. NASA-Earth Observatory (NASA-EO). Measuring vegetation (NDVI & EVI). Available at [http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring\\_vegetation\\_2.php](http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php) (accessed online 16 December 2014).
14. Aitchison, J. *The Statistical Analysis of Compositional Data*; Chapman and Hall: London, UK, 1986.
15. Aitchison, J.; Greenacre, M. Biplots of compositional data. *Appl. Stat.* **2009**, *51*, 375–392.
16. Filzmoser, P.; Hron, K.; Reimann, C. Principal component analysis for compositional data with outliers. *Environmetrics* **2009**, *20*, 621–632.
17. Hron, K.; Jelinkov, M.; Filzmoser, P.; Kreuziger, R.; Bednar, P.; Bartak, P. Statistical analysis of wines using a robust compositional biplot. *Talanta* **2012**, *90*, 46–50.
18. Egozcue, J.J.; Pawlowsky-Glahn, V.; Mateu-Figueraz, G.; Barceló-Vidal, C. Isometric logratio transformations for compositional data analysis. *Math. Geosci.* **2003**, *35*, 279–300.
19. Filzmoser, P.; Hron, K. Outlier detection for compositional data using robust methods. *Math. Geosci.* **2008**, *40*, 233–248.

20. Funk, C. A Climate Trend Analysis of Kenya—August 2010. Available online: <http://pubs.usgs.gov/fs/2010/3074/> (accessed on 12 June 2014).
21. Collier, P.; Conway, G.; Venables, T. Climate change and Africa. *Oxf. Rev. Econ. Policy* **2008**, *24*, 337–353.
22. Hope, K.R. Climate change and poverty in Africa. *Int. J. Sustain. Dev. World Ecol.* **2009**, *16*, 451–461.
23. Cooper, P.; Dimes, J.; Rao, K.; Shapiro, B.; Shiferaw, B.; Twomlow, S. Coping better with current climatic variability in the rain-fed arming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? *Agric. Ecosyst. Environ.* **2008**, *126*, 24–35.
24. Durack, P.J.; Wijffels, S.E.; Matear, R.J. Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science* **2012**, *336*, 455–458.
25. Nicholson, S.E. The nature of rainfall variability over Africa on time scales of decades to millenia. *Glob. Planet Chang.* **2000**, *26*, 137–158.
26. Morton, J.; Barton, D. Destocking as a Drought Mitigation Measure: Clarifying rationales and answering critiques. *Disasters.* **2002**, *26*, 213–228.
27. Thornton, P.K.; Jones, P.G.; Ericksen, P.J.; Challinor, A.J. Agriculture and food systems in sub-Saharan Africa in a 4 °C+ world. *Philos. Trans. R. Soc. A* **2011**, *369*, 117–136.
28. Muller, C.; Cramer, W.; Hare, W.L.; Lotze-Campen, H. Climate change risks for African agriculture. *Proc. Natl. Acad. Sci. USA* **2010**, *108*, 4313–4315.
29. López-Carr, D.; Pricope, N.G.; Aukema, J.E.; Jankowska, M.M.; Funk, C.; Husak, G.; Michaelsen, J. A spatial analysis of population dynamics and climate change in Africa: Potential vulnerability hot spots emerge where precipitation declines and demographic pressures coincide. *Popul. Environ.* **2014**, *35*, 323–339.
30. Justus, F.; Yu, D. Spatial distribution of greenhouse commercial horticulture in Kenya and the role of demographic, infrastructure and topo-edaphic factors. *ISPRS Int. J. Geo-Inf.* **2014**, *3*, 274–296.
31. Ulrich, A.; Speranza, C.I.; Roden, P.; Kiteme, B.; Wiesmann, U.; Nüsser, M. Small-scale farming in semi-arid areas: Livelihood dynamics between 1997 and 2010 in Laikipia, Kenya. *J. Rur. Stud.* **2012**, *28*, 241–251.
32. Sayer, J.; Cassman, K.G. Agricultural innovation to protect the environment. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8345–8348.
33. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20260–20264.
34. Brondizio, E.S.; Moran, E.F. Human dimensions of climate change: The vulnerability of small farmers in the Amazon. *Philos. Trans. R. Soc. B* **2008**, *363*, 1803–1809.
35. Kareri, R. Some aspects of the geography of Kenya. In Proceedings of the 2010 Indiana University-Perdue University Indianapolis (IUPUI) Fulbright Hays Group Projects Symposium, Eldoret, Kenya, 9 July–3 August 2010.
36. Notter, B. Rainfall-Runoff Modeling of Meso-Scale Catchments in the Upper Ewaso Ng’iro Basin, Kenya. Master’s Thesis, University of Berne, Bern, Switzerland, 2003.
37. Beck, H.E.; McVicar, T.R.; van Dijk, A.I.J.M.; Schellekens, J.; de Jeu, R.A.M.; Bruijnzeel, L.A. Global evaluation of four AVHRRNDVI data sets: Intercomparison and assessment against Landsat imagery. *Remote Sens. Environ.* **2011**, *115*, 2547–2563.

38. Slayback, D.A.; Pinzon, J.E.; Los, S.O.; Tucker, C.J. Northern hemisphere photosynthetic trends 1982–99. *Glob. Chang. Biol.* **2013**, *9*, 1–15.
39. Samanta, A.; Ganguly, S.; Hashimoto, H.; Devadiga, S.; Vermote, E.; Knyazikhin, Y.; Nemani, R.R.; Myneni, R.B. Amazon forests did not green-up during the 2005 drought. *Geophys. Res. Lett.* **2010**, *37*, L05401.
40. World Resources Institute. Download Kenya GIS Data. Available online: <http://www.wri.org> (accessed on 12 September 2014).
41. Albaladejo, J.P.; Fernandez, J.M.; Soto, J.A. Dealing with distances and transformations for fuzzy C-means clustering of compositional data. *J. Classif.* **2012**, *29*, 144–169.
42. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. *Remote Sens. Rev.* **1995**, *13*, 95–120.
43. Thió-Henestrosa, S.; Martín-Fernández, J.A. Detailed guide to CoDaPack: A freeware compositional software. In *Compositional Data Analysis in the Geosciences: From Theory to Practice*; Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V., Eds.; Geological Society: London, UK, 2006; Volume 264, pp. 101–118.
44. Templ, M.; Hron, K.; Filzmoser, P. Robcompositions: Robust Estimation for Compositional Data. Manual and Package, Version 1.3.3. Available online <http://cran.r-project.org/package=robCompositions> (accessed on 16 December 2014).
45. Smith, P.G. Automated log-ratio analysis of compositional data: Software suited to analysis of habitat preference from radiotracking data. *Bat Res. News* **2004**, *45*, 16.
46. Smith, P.G. *Compos Analysis Version 6.3 User's Guide*; Smith Ecology Ltd.: Abergavenny, UK, 2010.
47. Jessica, E.; Tierney, J.E.; Smerdon, K.J.; Anchukaitis, R.S. Multidecadal variability in East African hydroclimate controlled by the Indian Ocean. *Nature* **2013**, *493*, doi:10.1038/nature11785.
48. Williams, C.A.; Hanan, N.P. ENSO and IOD teleconnections for African ecosystems: Evidence of destructive interference between climate oscillations. *Biogeosciences* **2011**, *8*, 27–40.
49. Nicholson, S.E.; Davenport, M.L.; Malo, A.R. A comparison of the vegetation response to rainfall in the Sahel and East Africa, using normalized difference vegetation index from NOAA AVHRR. *Clim. Chang.* **1990**, *17*, 209–241
50. Anyamba, A.; Tucker, C.J.; Eastman, J.R. NDVI anomaly patterns over Africa during the 1997/98 ENSO warm event. *Int. J. Remote Sens.* **2001**, *22*, 1847–1859.
51. Luo, J.J.; Zhang, R.; Behera, S.K.; Masumoto, Y.; Jin, F.F.; Lukas, R.; Yamagata, T. Interaction between El Niño and extreme Indian Ocean Dipole. *J. Clim.* **2010**, *23*, 726–742.
52. Huho, J.M.; Mugalavai, E.M. The effects of droughts on food security in Kenya. *Int. J. Clim. Chang.: Impacts Resp.* **2010**, *2*, 61–72.
53. International Development Research Center (IDRC). *Reducing Vulnerability among Pastoralists in Northern Kenya. Climate Change Adaptation in Africa Program 2009–10 Annual Report*; IDRC: Ottawa, ON, Canada, 2010.
54. Longley, C.; Wekesa, M. Improving Drought Response in Pastoral Areas of Kenya; Lessons and recommendations. Available online: <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinionfiles/5299.pdf> (accessed on 25 September 2014).
55. Shisanya, C.A. The 1983–1984 drought in Kenya. *J. East. Afric. Res. Dev.* **1990**, *20*, 127–148.

56. Williams, C.A.; Hanan, N.P.; Baker, I.; Collatz, G.J.; Berry, J.; Denning, A.S. Interannual variability of photosynthesis across Africa and its attribution. *J. Geophys. Res.* **2008**, *113*, G04015.
57. Redmon, L.A. Effects of Drought on Plant Growth. Available online: [http://hays.agrilife.org/files/2011/08/May20111\\_3.pdf](http://hays.agrilife.org/files/2011/08/May20111_3.pdf) (accessed online 25 September 2014).
58. Grünewald, F.; Robins, K.; Odicoh, A.; Nicholson, N. Real Time Evaluation of the Drought Response in the Horn of Africa. Available at <http://www.alnap.org/resource/3444.aspx> (Accessed on 7 January 2015)
59. Annamalai, H.; Xie, S.P.; McCreary, J.P.; Murtugudde, R. Impact of Indian Ocean sea surface temperature on developing El Niño. *J. Clim.* **2005**, *18*, 302–319.
60. Anyamba, A.; Eastman, J.R. Interannual variability of NDVI over Africa and its relation to El Niño Southern Oscillation. *Int. J. Remote Sens.* **1996**, *17*, 2533–2548.
61. McNaughton, S.J.; Sala, O.E.; Oesterheld, M. Ecosystem-level pattern of primary productivity and herbivory in terrestrial habitats. *Nature* **1989**, *341*, 142–144.
62. Lyon, B.; DeWitt, D. A recent and abrupt decline in the East African long rains. *Geophys. Res. Lett.* **2012**, *39*, L02702.
63. Nicholson, S.E.; Kim, J. The relationship of the El Niño Southern Oscillation to African rainfall. *Int. J. Climatol.* **1997**, *17*, 117–135.

© 2015 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 license (<http://creativecommons.org/licenses/by/4.0/>).