Environmental Development 28 (2018) 55–66

Contents lists available at ScienceDirect

Environmental Development

journal homepage: www.elsevier.com/locate/envdev

Toward an integrated watershed zoning framework based on the spatio-temporal variability of land-cover and climate: Application in the Volta river basin

Dagbegnon Clement Sohoulande Djebou

USDA-ARS Coastal Plain Soil, Water and Plant Conservation Research Center, 2611 West Lucas Street, Florence, SC 29501, USA

ARTICLE INFO

Keywords:
Watershed management
Entropy theory
Cluster analysis
Leaf area index
Ecosystem
Climate variability
Volta river basin

ABSTRACT

This article details a watershed regionalization approach which uses the concept of entropy in combination with the k-means clusters analysis. The regionalization approach aims to subdivide a watershed into relatively uniform zones based on the spatial variability of the local climate and land-covers. A case study is presented to illustrate the approach and outline the environmental implications of the outcomes. Especially, the study reports an application in the Volta river basin which is a transnational watershed, shared by six different countries in West Africa. Over years, the transboundary status of the Volta watershed seems to have exacerbated its environmental challenges, because the environmental policies in the six countries do not necessarily complement. Subsequently, it is desirable to envision unified scientific tools to support the management platform of the Volta basin. To date, the literature on the Volta has virtually neglected this aspect. Hence, this case study is timely as it intends to create a unified zoning system for the Volta river basin. In the study, formulations of entropy theory and k-means clustering were jointly applied to 16-years gridded time-series of monthly leaf area index, precipitation, and temperature across the Volta basin. Based on a clustering optimization criterion, a total of five zones were identified then the related land-cover and climatic patterns were comparatively analyzed. Significant environmental contrasts were diagnosed then specificities were pinpointed for each zone. A comparison of the new zones with an existing macro-scale ecoregion shows similarities which sustain the capacity of the regionalization approach to capturing meaningful biophysical signals. Hence, the zoning technique may be valued for further applications in environmental management.

1. Introduction

The climatic diversity of the globe plays an important role in controlling the natural equilibrium of terrestrial ecosystems (Pecl et al., 2017; Sala et al., 2000). For that reason, climate variability is frequently used in combination with other biophysical factors (e.g. soil, land-cover, agricultural production) to identify ecological regions (Fischer et al., 2002; Bailey, 1989). Over years, researchers have developed macroscale ecoregions at continental and global levels (Abell et al., 2008; Bailey and Hogg, 1986). However, the macroscale regions do not often emphasize specificities needed for implementing environmental management strategies at local levels. As a result, studies are frequently conducted to generate or update ecological regions within specific administrative domains (i.e. country, state). In such a circumstance, the regionalization frameworks do not necessarily display a consistent transboundary continuity. Moreover, when natural resources span over multiple administrative territories, these resources are likely

E-mail address: Clement.sohoulande@ars.usda.gov.

https://doi.org/10.1016/j.envdev.2018.09.006
Received 12 November 2017; Received in revised form 24 September 2018; Accepted 24 September 2018
2211-4645/ Published by Elsevier B.V.

subjected to different management policies which may not complement from an environmental perspective (Jägerskog et al., 2012). This lack of complementarity is often reported with a wide range of biophysical implications which are perceptible at the watershed scale especially when the watershed encompasses multiple administrative territories (Jägerskog et al., 2012). Indeed, many transboundary watersheds face serious environmental and ecosystem issues due to the lack of a coherent management agenda. Thenceforth, it is rational to envision more concordant environmental management plans for regional watersheds which are often shared by different communities or countries (Hornbeck and Swank, 1992).

At the watershed level, the diversity of local ecosystems is often influenced by factors such as the climate, the hydrography, the soil and the extent of human actions (Sala et al., 2000). These factors sustain multiple interactions which determine the biophysical functionality of the watershed. In addition to their hydrological functions, watersheds are hubs for fauna and flora communities which interact mutually (Hornbeck and Swank, 1992). The watershed also shelters human communities, but in opposition to the others biological communities, human factors are generally recognized to be causing serious ecological disturbances across watersheds. For instance, it is demonstrated that several watersheds are subjected to an alteration of their hydrological components because of the growth of human population and the inherent pressure on natural resources (Sala et al., 2000). Yet, the magnitude of ecosystems degradation often varies from one location to another depending on the demography and the economic activities across a watershed. However, a change that occurs in a given location is likely to have a direct or indirect impact on other locations of the watershed. This is true because the functionality of a watershed generally engages multiple biophysical components which interplay diversely in time and space (Castro et al., 2016; Hornbeck and Swank, 1992). Thus, preserving a healthy watershed is part of the complex equation of ecosystem sustainability, but its achievement may require a reconsideration of the watershed as a functional unit.

To date, the popular maps used in environmental management (i.e. climatic maps, agro-ecological maps) have preferentially emphasized administrative boundaries rather than the watershed. This reality may rapidly change in the future as the ongoing challenges related to the environmental sustainability may compel natural resources managers to emphasize watersheds. However, there are no well-established watershed zoning approaches which use multiple biophysical factors at the same time. In regard to this shortcoming, studies are needed to develop watershed zoning techniques capable to capture tangible climatic and biological variabilities. The content of this study concurs with this idea as it expounds a multivariate watershed zoning technique which may be used as an environmental management tool.

In general, univariate spatial zoning methods are straightforward and they more likely involve one or two dimensions (i.e. temporal and spatial). However, the zoning task becomes very complicated when multiple variables need to be taken into consideration at the same time (Fischer et al., 2002). This is the case with the zoning framework presented in this paper. Indeed, the study describes a zoning approach which is elaborated to capture simultaneously the bio-climatic variabilities of watersheds. The theoretical approach combines the concept of entropy with the k-means cluster analysis (Singh, 2011; Yoon et al., 2007). This paper reports the zoning approach using a case study of the Volta river basin which is a transboundary watershed located in West Africa (Rodgers et al., 2006; van de Giesen et al., 2001). The biophysical variables involved in the study include precipitation, temperature and leaf area index (LAI). The outcomes of the studies are presented by this paper which is structured into six sections including: (i) this introduction, (ii) the study region, (iii) the data and method, (iv) the results, (v) the discussion, and (vi) the conclusion.

2. Study region

This study addressed the Volta river basin, a watershed located in West Africa (Fig. 1). The Volta basin covers approximatively 385,000 km^2 and spans between the longitudes 2.25E and 4.75W, the latitudes 5.75N and 14.75N. Characterized, by a shallow watercourse, the Volta is a transboundary watershed, shared by six countries including Benin (4.1% of the basin which equals 14.2% of Benin's territory), Burkina Faso (46.4% of the basin which equals 66.8% of Burkina Faso's territory), Cote d'Ivoire (1.8% of the basin which equals 2.2% of Cote d'Ivoire's territory), Ghana (38.6% of the basin which equals 63.7% of Ghana's territory), Mali (2.4% of the basin which equals 0.8% of Mali's territory), and Togo (6.7% of the basin which equals 45.4% of Togo's territory) (Bharati et al., 2008). Fig. 2 portrays the transboundary status of the watershed. Independently, these six countries had developed their own agro-climatic zones' maps, but the resulting zones do not necessarily extend continuously beyond administrative boundaries. As an example, Antwi et al. (2014) reported an ecoregion subdivision of Ghana which seems very detailed but does not display a direct continuity with zonings developed for Cote d'Ivoire which is a neighboring country (Brou et al., 2005). In this circumstance, it is difficult to retrieve a unique zoning system for the Volta watershed. This is the rationale behind the choice of the Volta river basin in this study, then the outcomes may be considered as a supporting tool for its ongoing environmental management strategies.

3. Data and method

3.1. Data

The objective of the zoning approach in this study is to determine regions with high similarity based on the spatial and temporal patterns of multiple biophysical variables. Note, a common fact in multivariate analytical frames is that the candidate variables are not necessarily equally relevant. For this reason, it is often encouraged to identify and retain pertinent variables. In the case of this research, three biophysical variables are considered including leaf area index (LAI), precipitation (P) and surface air temperature (T). These three variables are $0.5^\circ \times 0.5^\circ$ gridded monthly time-series spanning the period February 2000 to December 2015. Note, data availability is the main justification for the targeted time period. A total of 174 grids evenly distributed, was selected within the

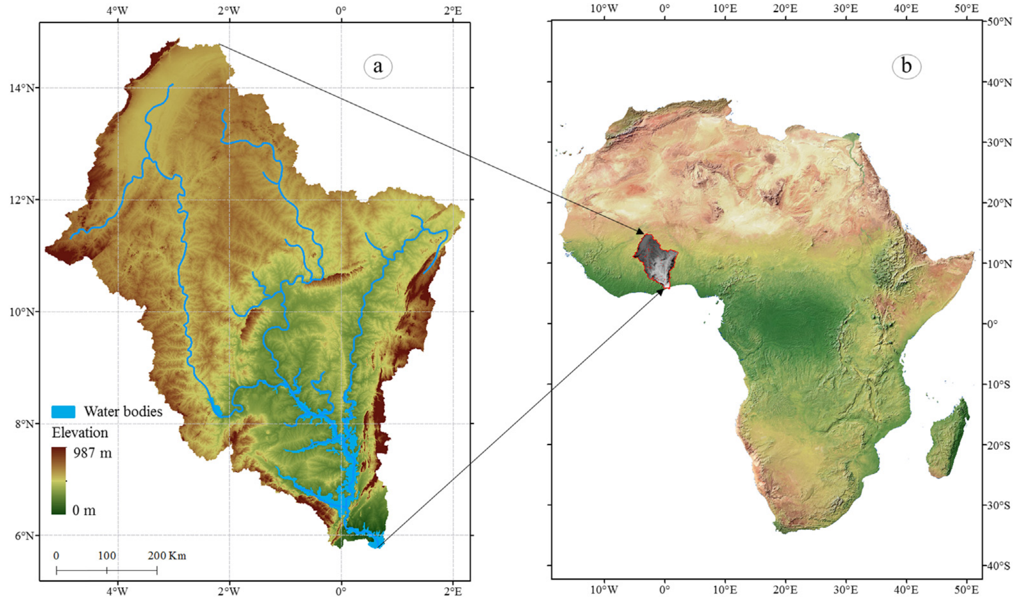

Fig. 1. The Volta river basin as located in the African continent. (a) Presents the topography contrast across the watershed as well as the hydrography. (b) Spots the Volta watershed on Africa map.

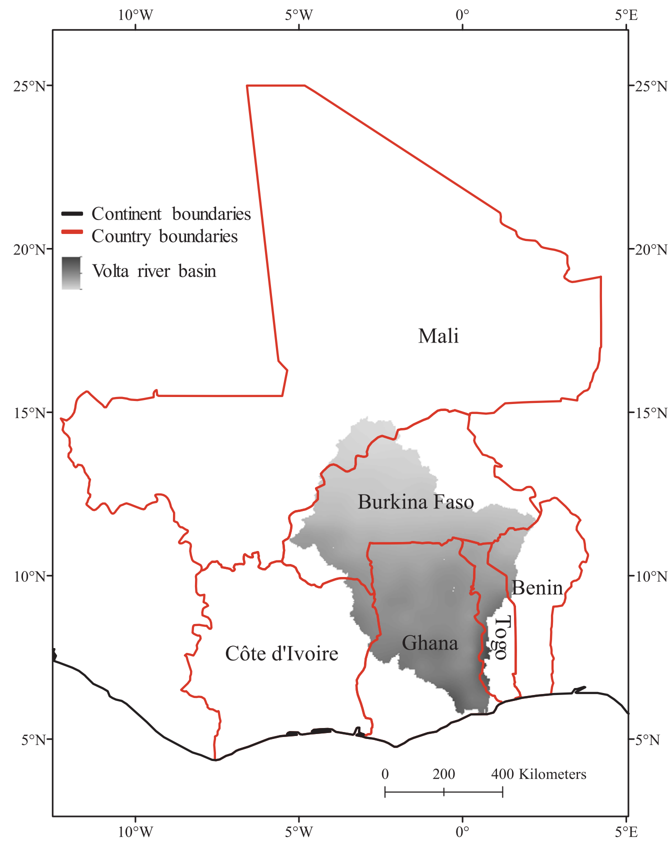

Fig. 2. Transboundary status of the Volta river basin.

spatial boundaries of the Volta watershed (Fig. 3a). Each of the grids is associated with time-series of the three variables (i.e. LAI, P and T). The principle of the zoning technique is to group the grids in an optimal number of spatial subsets, such that each subset will correspond to a unique zone. Individually, the three variables included in the study are pertinent and concordant with the scope of

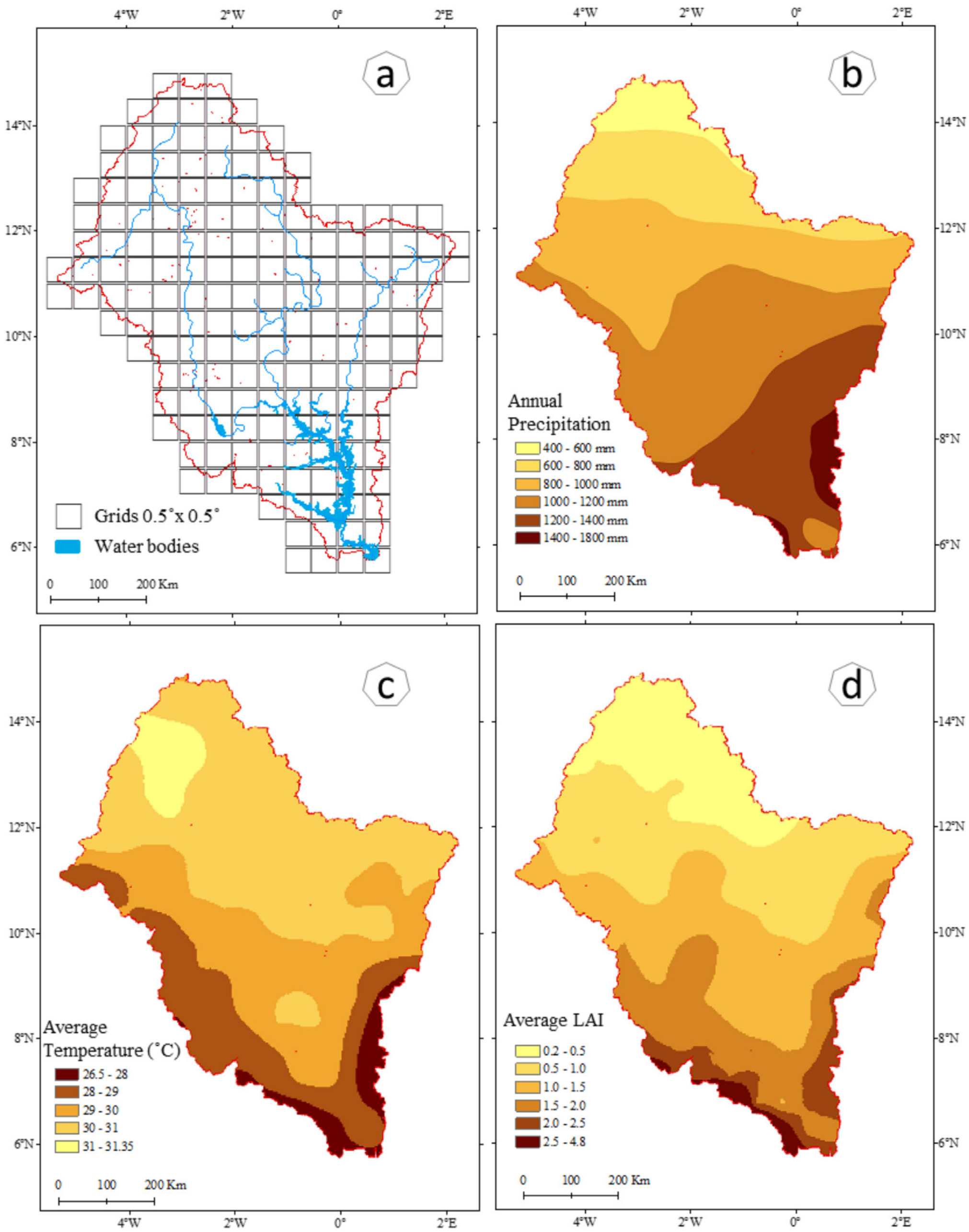

Fig. 3. Gradients of annual precipitation (b), average temperature (c) and average LAI across the Volta River basin. Display of the 174 grids (0.5° × 0.5°) encompassed within the watershed boundaries (a).

the study. Indeed, precipitation and temperature are frequently used in climatic regionalization while leaf area index is a critical indicator of vegetation canopy and structural properties (Almazroui et al., 2015; Garrigues et al., 2008). Fig. 3b, c, and d expound respectively the gradient of precipitation, temperature, and LAI across the Volta river basin.

Worldwide, LAI is preferentially utilized to assess land-cover change in time and space (Huete et al., 2010; Garrigues et al., 2008). LAI is a measurement of the structural property of the vegetation canopy. In practice, LAI is defined as the equivalent of the above ground plants leaves displayed over a unit surface (Zhang et al., 2008). Aligning with this definition, statistics of LAI (e.g. average, ranges) are frequently used to quantify land-cover change as it is in this study. With the advances in remote sensing, the use of LAI is found practical for large-scale land-cover assessment. For instance, several authors successfully used LAI to interpret trends of agriculture, deforestation, or urbanization sprawl at the watershed scale (Zhu et al., 2016; Xavier and Vettorazzi, 2004). The LAI data employed in the study are developed by the National Aeronautics and Space Administration's Earth Observations (NASA-NEO) (www.neo.sci.gsfc.nasa.gov). The NASA-NEO's LAI data are derived from the Moderate Resolution Imaging Spectroradiometer

(MODIS) (Knyazikhin et al., 1999). However, these LAI data are available only from February 2000. In order to comply with the temporal availability of both LAI and the climate variables (P and T), the study addressed the period February 2000 to December 2015 (191 months). Yet, the precipitation and temperature data used in the study are obtained from the $0.5^{\circ} \times 0.5^{\circ}$ gridded time-series dataset produced by the University of East Anglia's Climatic Research Unit (CRU) and released by the Centre for Environmental Data Analysis (CEDA) (www.ceda.ac.uk). Particularly the CRU's time-series version 4.00 dataset (Harris and Jones, 2017) was considered, then monthly precipitation and temperature data were retrieved distinctly for the 174 grids encompassed by the Volta watershed.

3.2. Method

The literature reports various techniques often used for spatial regionalization. The well-known techniques are probably those based on principal components or factors analysis (Santos et al., 2017; Almazroui et al., 2015). Authors also used cluster analysis as a spatial regionalization technique with satisfactory results (Zhang et al., 2016). In either case, the regionalization principle consists of grouping spatial elements (i.e. spots, locations, stations or grids) into subsets with similar characteristics. Often, these elements are associated with time-series of a given biophysical variable. Hence, these regionalization techniques only apply to one biophysical factor at the time. For instance, the spatial regionalization technique using principal components is often applied when a single biophysical factor is considered (Demšar et al., 2013). Subsequently, the resulting zones are typically the expression of the initial biophysical factor. In order to address multiple biophysical factors at a time, a two-stage pooling procedure is more likely needed. The theoretical approach reported in the present article is an illustration of the two-stage pooling procedure.

The study approach is based on a combination of the entropy theory and the k-means cluster analysis. Explicitly, the analytical procedure consists of two stages. The first stage addresses the temporal variability of the targeted variables. At this stage, the entropy theory is applied separately to the gridded time-series of LAI, P, and T. Also known as information theory or Shannon Entropy (Shannon, 1948), the concept of entropy is a mathematic formulation developed for quantifying disorder (Singh, 2011). The entropy theory is well-established for variability assessment in water resources and environmental research (Singh, 2013). This study employs the concept of entropy with the perspective of capturing the essential variability contained in time-series data. Especially, the discrete formulation of Shannon entropy was considered then used for the computation of the entropy values of LAI, P, and T. The computations were conducted distinctly for each of the 174 grids encompassed within the spatial boundary of the Volta watershed. Note, the discrete formulation of the Shannon entropy uses discrete probability distributions (Singh, 2013). The discrete probabilities were estimated based on a relative frequency analysis. Especially, for each grid, *N* equal-range bins were distinctly drawn from the time-series of LAI, P, and T. The integer *N* (i.e. the number of bins) was identified using the formulation proposed by Sturges (1926) which is given by the Eq. (1).

$$N = 1 + 3.3 * Log \ (Sample \ \ Size) \tag{1}$$

In the study, the size of the monthly time-series is 191. Subsequently, the number *N* of bins considered in the entropy computation is $N = 9$. Given a variable x and a related bin x_i, the corresponding discrete probability $p(x_i)$ is drawn such that the Eq. (2) is verified. Implicitly, the entropy $H(x)$ is calculated using the Eq. (3).

$$\sum_{i=1}^{N} p \ (x_i) = 1 \tag{2}$$

$$H(x) = -\sum_{i=1}^{N} p \ (x_i) \ \log_2[p \ (x_i)] \tag{3}$$

The entropy values were estimated for the three variables independently, such that each grid was linked with three scores corresponding respectively for the time-series of LAI, P and T. The second stage of the theoretical approach employed a k-means cluster analysis (Yoon et al., 2007). In opposition to the entropy application which addressed the 174 grids separately, the k-means cluster analysis involved all the grids at a time. Note, each grid is represented by three scores which are the corresponding entropy estimates of LAI, P, and T. The goal of the k-means clustering is to generate k subsets or centroids such that the Euclidean distance D is minimized (Yoon et al., 2007). The estimate of D uses the Eq. (4) where k stands for the number of clusters, S_j designates any subset or cluster such that $j \in \{1, \ 2, \ \ldots, \ k\}$. μ_j is the mean associated with the centroid S_j, and y_l is the score of any element $l \in S_j$.

$$D = \sum_{j=1}^{k} \sum_{l \in S_j} |y_l - \mu_j|^2 \tag{4}$$

The k-means clustering procedure is iterative (Yoon et al., 2007). During the clustering process, elements located in the spatial domain of the watershed (i.e. grids) are individually assigned to one of the k subsets. Initially, the number k of clusters is arbitrary fixed and it is an integer which can range from 1 to the sample size (i.e. $1 \le k \le 174$). Explicitly, all grids can be agglomerated in a single group or successively divided into smaller groups till each grid is assimilated with a separate cluster. However, for a spatial regionalization study, it is prominent to identify an optimum k which would correspond to the number of zones (Milligan and Cooper, 1985). Interestingly, k can be optimally identified through an iterative process where the within-cluster homogeneity is evaluated based on the cubic clustering criterion (CCC) (Ketchen and Shook, 1996; Milligan and Cooper, 1985). In fact, the cubic clustering criterion is a fit statistic which is relevant for identifying the optimum value of k during a k-means clustering (Yoon et al., 2007). The

Table 1
Identifying optimal number of clusters using the CCC criterion in a k-means clustering analysis. (*) indicates the maximum CCC values which corresponds to the optimum numbers of clusters.

Number of clusters	CCC	Distribution of the 174 grids within the clusters
2	−0.05	115; 59
3	−0.08	109; 57; 8
4	1.39	97; 46; 23; 8
5	2.72*	54; 45; 42; 23; 10
6	2.04	56; 51; 35; 23; 5; 4
7	0.38	55; 51; 36; 23; 4; 4; 1
8	2.41	39; 39; 30; 30; 19; 8; 5; 4

process consists of performing the clustering analysis using different values for k. For instance, the present study used consecutive integers ranging from 2 to 8. For each value of k, the k-means clustering is conducted and the corresponding CCC is evaluated. Ultimately, the optimum number of cluster is the k associated with the peak of CCC (Ketchen and Shook, 1996). The detailed outcomes of these analyses are now provided in the next section of this article.

4. Results

4.1. Watershed zoning framework

This section reports the outcomes of the application of the zoning approach in the Volta watershed. The entropy values were estimated, thereafter the k-means cluster analysis was conducted according to the methodology described in Section 3.2. Table 1 summarizes the k-means clustering performed on the multivariate entropy scores. Based on the clustering criterion CCC analysis (Milligan and Cooper, 1985), the optimum number of clusters was identified as k = 5 (Table 1). Thus, the k-means clustering analysis yielded 5 distinct clusters. Each cluster is a group of grids identified individually by their geographic coordinates (i.e. longitude and latitude). As a result, the spatial domains associated with the five subsets were projected and presented in Fig. 4a. Indeed, Fig. 4a illustrates the spatial locations of the elements of the five clusters within the boundary of the Volta watershed. Notice that in Fig. 4a, the elements pertaining to each subset span continuously within the same spatial domain. The spatial continuity observed here is very important as it portends the performance of the regionalization techniques. It also allows to agglomerate the grids for the zoning. Hence, zones were delineated based on the clusters and the resulting zoning framework is presented in Fig. 4b. From a visual prospect, the delineated zones show tangible disparities in term of shape, location, and areas. Approximatively, zone 1 covers 2.8% of the watershed, zone 2 covers 28.8%, zone 3 covers 31.9%, zone 4 covers 24.2% and zone 5 covers 12.3%. These physical disparities

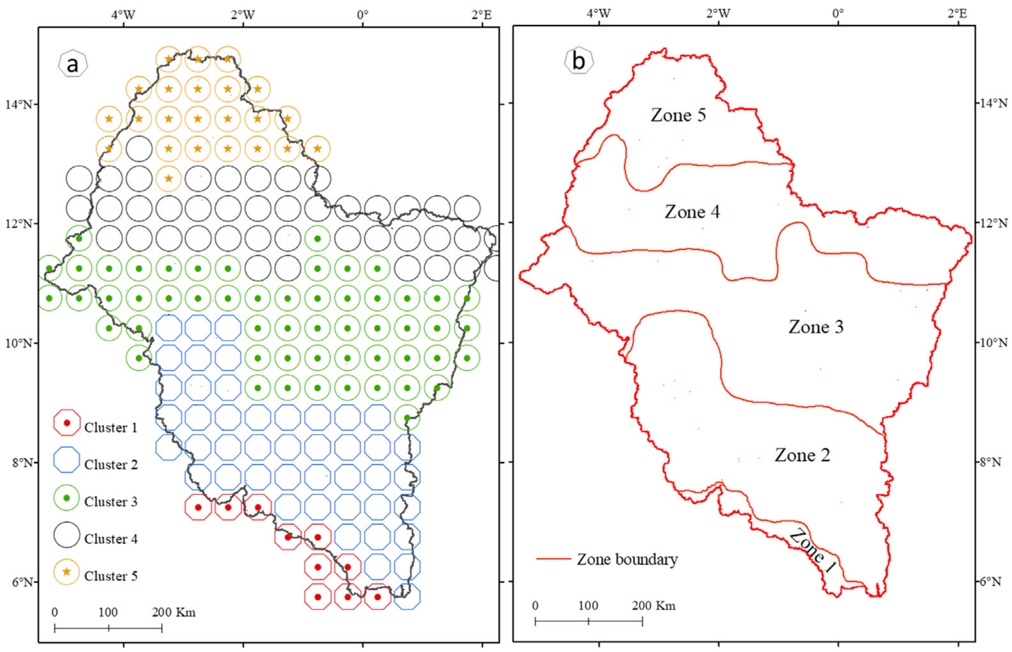

Fig. 4. Zoning of the Volta watershed based on the combined k-means cluster and entropy theory assessment of precipitation, temperature and leaf area index LAI. Note, zone 1 corresponds to cluster 1, zone 2 to cluster 2, zone 3 to cluster 3, zone 4 to cluster 4, and zone 5 to cluster 5.

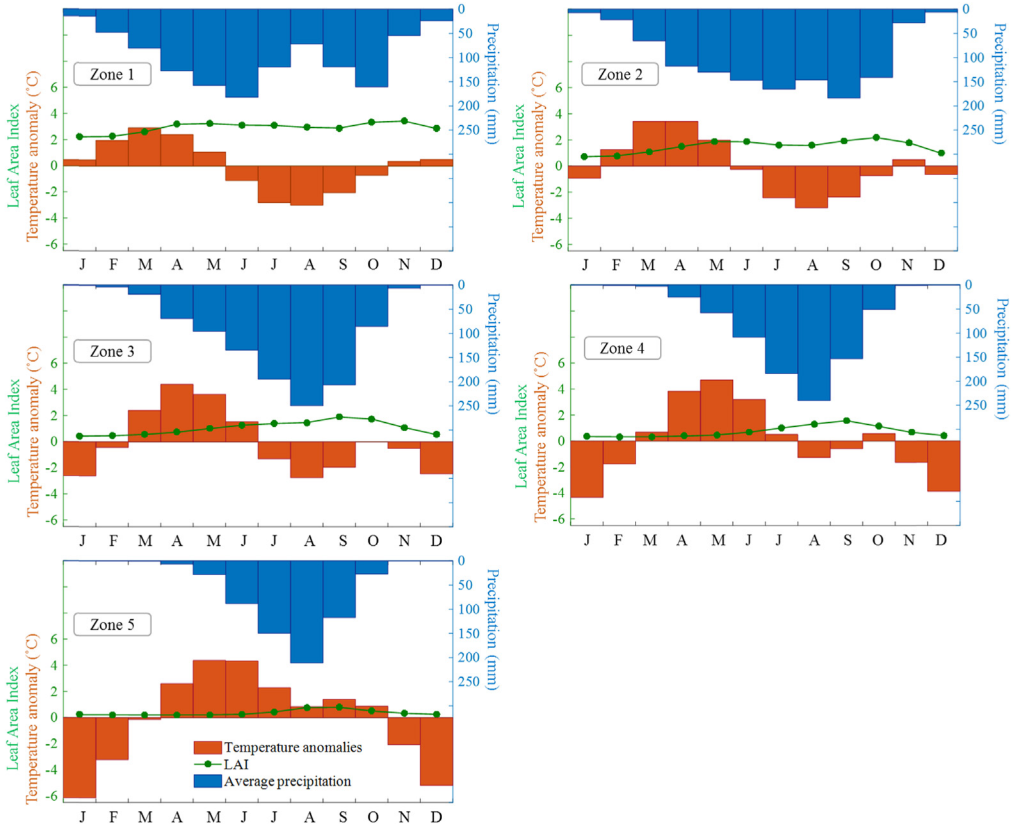

Fig. 5. Monthly patterns of precipitation, temperature and LAI across the Volta river basin. The averages are estimated for individual zone using time-series data of the period February 2000 to December 2015. The climograph of precipitation report the average monthly total precipitation while the climograph of temperature reports the average monthly anomalies. Note, the abscissa of each figure is labeled with the initial of the successive months of the year (J = January, F = February, …, D = December).

portend crucial biological and climatic specificities which will be elucidated in the next sections of this article. When the variables are taken individually, it is difficult to establish a clear similitude between the delineated zones in Fig. 4b and the gradients of P (Fig. 3b), T (Fig. 3c) and LAI (Fig. 3d). This observation is perhaps an indication that the zoning framework developed is not driven by only one biophysical factor (i.e. LAI or P or T), instead, it is the resultant of all the three variables together. Nevertheless, the spatial domain corresponding to each zone is in a single extension, but this continuity seems to align with the overall latitudinal gradient previously observed with LAI, P, and T (Fig. 3b, c, d). The comparative analysis of the land-cover and climatic specificities of the zoning framework is detailed in the next section.

4.2. Biophysical facets of the zones framework

This section examines and compares the land-cover and climatic specificities of the delineated zones. Fig. 5 portrays the seasonal patterns of precipitation, temperature, and LAI across the zones. Precisely, the graphs in Fig. 5 are drawn based on the average monthly values, estimated separately for the individual months of the year (i.e. January, February, …, December). However, the climograph of temperature presents the monthly anomaly which is the deviation from the long-term average temperature. Table 2 complements Fig. 5 by diagnosing and recapping the seasonality of land-cover and climate features in the five zones. The comments on land-cover greenness reported in Table 2, align with Zhu et al. (2016) who recently addressed the earth greenness by referring to LAI values. The average values of LAI commented in Table 2 are sustained by the statistical analysis reported in Table 3. Especially, the analysis of Table 3 shows a gradual decrease of LAI and total precipitation from zone 1 to zone 5. In the same time, the average temperature increases gradually from zone 1 to zone 5. The climographs of precipitation presented in Fig. 5 reveal tangible signals of seasonality. Thus, the climograph of precipitation is bimodal in zone 1 but unimodal in zones 2, 3, 4 and 5. Table 3 presents a comparative analysis of LAI, precipitation and temperature base on a pairwise Student's *t*-test of means. The result of this statistical analysis corroborates with the contrasts summarized in Table 2. For instance, the pairwise comparison of the means with p-value = 0.05, revealed significant differences which allow confirming that the density of vegetation decreases from zone 1 to zone 5. Regarding precipitation and temperature, the pairwise Student's t statistics also revealed significant differences, but these differences seem more concordant with the overall latitudinal contrasts observed in Fig. 3. In complement with the pairwise *t*-test, a Man-Kendall monotonic trend analysis (Yue et al., 2002) is carried out separately on the time-series of annual precipitation, temperature, and LAI

Table 2
Comparative analysis of the contrast of land-cover and climate within the five zones identified across the Volta river basin.

Regions	Land-cover features	Climate patterns
Zone 1	1. Greener compared to all zones 2. Average monthly LAI ranges from 1.50 to 4.23 3. Coefficient of Variation of LAI = 0.16 4. With a month time-lag, precipitation explains 29% of the variance of LAI	1. Bimodal precipitation distribution with peaks during June and October 2. Annual precipitation > 1200 mm 3. Annually, 6 months receive individually at least 10% of the annual precipitation and a total of 51% is received from April to July, then 25% from September to October. 4. April is the warmest month (30.1 °C); August is the coolest (24.7 °C)
Zone 2	1. Less green compared to zone 1 but greener than zones 3, 4 and 5. 2. Average monthly LAI ranges from 0.54 to 2.43 3. Coefficient of Variation of LAI = 0.33 4. With a month time-lag, precipitation explains 66% of the variance of LAI	1. Flat precipitation distribution reaching a peak during September 2. Annual precipitation > 1200 mm 3. Annually, 7 months receive individually at least 10% of the annual precipitation and a total of 90% is continuously received from April to October. 4. April is the warmest month (32.5 °C); August is the coolest (25.9 °C)
Zone 3	1. Less green compared to zones 1 and 2, but greener than zones 4 and 5. 2. Average monthly LAI ranges from 0.39 to 2.13 3. Coefficient of Variation of LAI = 0.47 4. With a month time-lag, precipitation explains 86% of the variance of LAI	1. Unimodal precipitation distribution with a peak during August 2. Annual precipitation ≈1100 mm 3. Annually, 5 months receive at least 10% of the annual precipitation and a total of 84% is continuously received from May to September. 4. April is the warmest month (34.33 °C); August is the coolest (27.18 °C)
Zone 4	1. Less green compared to zones 1, 2 and 3, but greener than zone 5. 2. Average monthly LAI ranges from 0.28 to 1.68 3. Coefficient of Variation of LAI = 0.58 4. With a month time-lag, precipitation explains 92% of the variance of LAI	1. Unimodal precipitation distribution with a peak during August 2. Annual precipitation ≈800 mm 3. Annually, 4 months receive individually at least 10% of the annual precipitation and a total of 83% is continuously received from June to September. 4. May is the warmest month (35.26 °C); January is the coolest (26.23 °C)
Zone 5	1. Less green compared to all zones 2. Average monthly LAI ranges from 0.19 to 0.92 3. Coefficient of Variation of LAI = 0.57 4. With a month time-lag, precipitation explains 89% of the variance of LAI	1. Unimodal precipitation distribution with a peak during August 2. Annual precipitation ≈600 mm 3. Annually, 4 months receive individually at least 10% of the annual precipitation and a total of 90% is continuously received from June to September. 4. May is the warmest month (35.14 °C); January is the coolest (24.61 °C)

Note: The statistics sustaining the comparisons in this table are presented by Tables 3 and 4. The use of LAI values as a greening indicator for land-cover aligns with Zhu et al. (2016).

Table 3
Comparing the delineated zones of the Volta watershed zones using a pairwise Student's *t*-test for mean comparison and the Man-Kendall monotonic trend analysis. The Student's t mean comparison is evaluated based on monthly precipitation, temperature and LAI data; the trend analysis is performed based on the annual data. For the trend analysis, the null hypothesis H_0: "there is no trend" is tested for individual zone and the Kendall's τ values are reported.

Spatial domain	Area (%)	Pairwise mean comparison using *t*-test			Kendall's τ		
		Precipitation (mm)	Temperature (°C)	LAI	Precipitation	Temperature	LAI
Zone 1	2.8	101.83^{a}	27.68^{d}	2.95^{a}	0.23	0.28	0.22
Zone 2	28.8	101.93^{a}	29.06^{c}	1.49^{b}	0.23	0.25	0.31
Zone 3	31.9	93.75^{b}	29.94^{b}	1.06^{c}	0.24	0.23	−0.33
Zone 4	24.2	72.59^{c}	30.59^{a}	0.72^{d}	0.14	0.38*	−0.25
Zone 5	12.3	55.34^{d}	30.78^{a}	0.37^{e}	0.22	0.29	0.37*

Note: The connecting letters a, b, c, d used for the *t*-test indicate the ranking of the means such that a > b > c > d; the means with same letter are considered equal at p-value = 0.05. Kendall's τ values with (*) indicates that H_0 is rejected at p-value < 0.05.

of each zone. Precisely, the Kendall' τ statistics were estimated and reported in Table 3. Unlike the results of the *t*-test analysis, the outcomes of the trend analysis did not exhibit significant patterns, and this is perhaps due to the relatively short time frame used in the analysis (i.e. 16 years). Nevertheless, this shortcoming should not rule out the consistency of the patterns depicted in Fig. 5, Tables 2, 3.

The disparities observed within the zones may also reflect signals of interactions embedded in the zoning framework. A good illustration of these interactions is here provided by the statistical analysis of the land-cover response to precipitation across the Volta watershed. Indeed, the response of LAI to precipitation is evaluated for each of the five zones using a Pearson correlation analysis with different time-lags (e.g. 0, 1, 2, 3, 4 months). As a result, the maximum response is obtained when a month time-lag is considered between precipitation and LAI. Table 4 summarizes the time-lag analysis and reports the patterns of the 0–4 months lags for each of the five zones. The month time-lag aligns with previous studies and is consistent with the monthly temporal scale of the

Table 4
Analysis of the time-lag of vegetation response to precipitation across the five zones identified across the Volta river basin.

Time-lag	R^2 estimated between LAI and Precipitation				
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5
No lag	0.22	0.45	0.56	0.6	0.55
1-month lag	**0.29**	**0.66**	**0.86**	**0.92**	**0.89**
2-months lag	0.08	0.28	0.57	0.57	0.47
3-months lag	0.01	0.01	0.09	0.1	0.07
4-months lag	0.01	0.04	0.03	0.02	0.02

Note: Time-series of monthly LAI and precipitation are correlated using different time-lags. Optimal lag is highlighted in bold.

precipitation and LAI data (Sohoulande Djebou et al., 2015). Fig. 6 exemplifies the land-cover response to precipitation with a month time-lag. The comparison of the trends shows meaningful disparities between the zones. For instance, the vegetation responses to precipitation appear minimal for zone 1 (R^2 = 0.29) and zone 2 (R^2 = 0.66) while the responses are maximal for zones 3, 4 and 5 ($0.86 \le R^2 \le 0.92$). In accordance with the analysis reported in Table 2, one may infer a higher sensitivity of vegetation-cover to precipitation in the drier zones of the watershed (i.e. zones 3, 4, 5). This result can be further understood by referring to the structural properties of the dominant vegetation covers in each zone. Yet, this facet of the analysis is reported in the discussion section which also examines the relevance of the zoning framework.

4.3. Comparing the new Volta zones with the existing world ecoregions

The delineated zones for the Volta basin are here compared with the United Nations Environment Program (UNEP)'s ecoregions also known as Bailey's world ecoregions (Bailey, 1989). The UNEP's ecoregions are the result of an international project developed during the 1980′s with the objective of subdividing the world into relatively uniform macro-ecoregions (Bailey and Hogg, 1986). The world ecoregions data used in this section are obtained from the United Nations Environment World Conservation Monitoring Centre datasets (www.unep-wcmc.org). Fig. 7a presents the UNEP's ecoregions comparatively to the Volta zoning framework. As highlighted by Fig. 7b, the ecoregions encompassing the Volta basing are: (i) "Mixed forests" described as mixed forests with short dry season; (ii) "Savannah forests" described as humid tall-grass savannahs and savannah forests; (iii) "Grassy savannahs" described as moderately

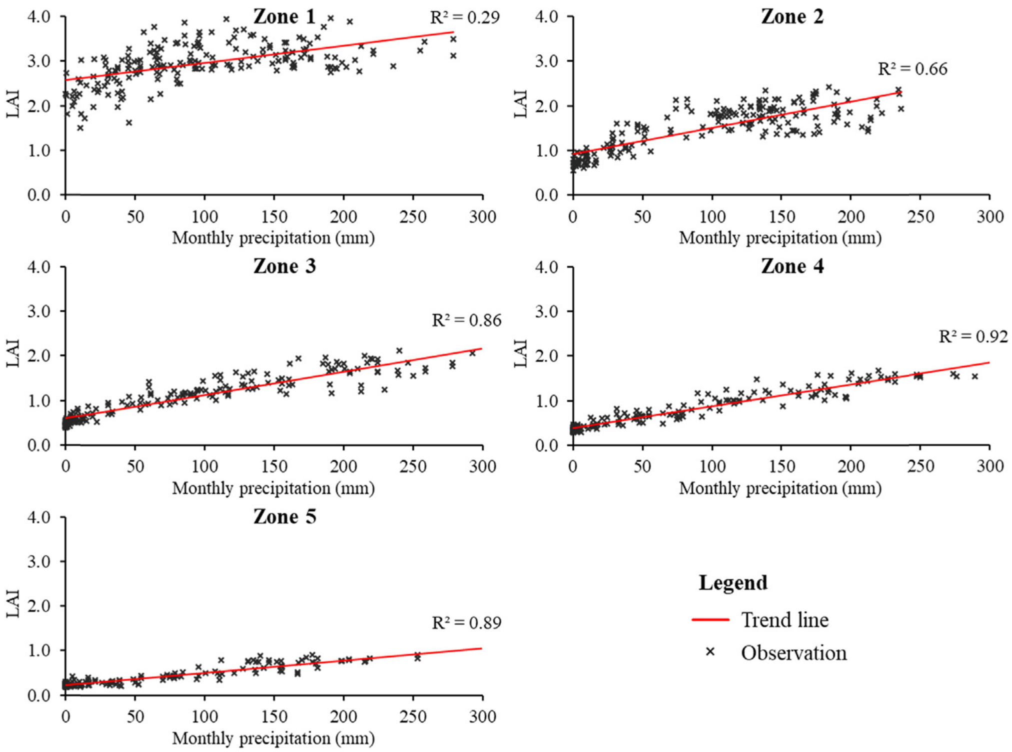

Fig. 6. Comparing vegetation sensitivity to precipitation in the zones delineated. A month time-lag is considered between monthly precipitation and monthly LAI values. The observations correspond to the average of the grids encompassed by each zone during the consecutive months from February 2000 to December 2015.

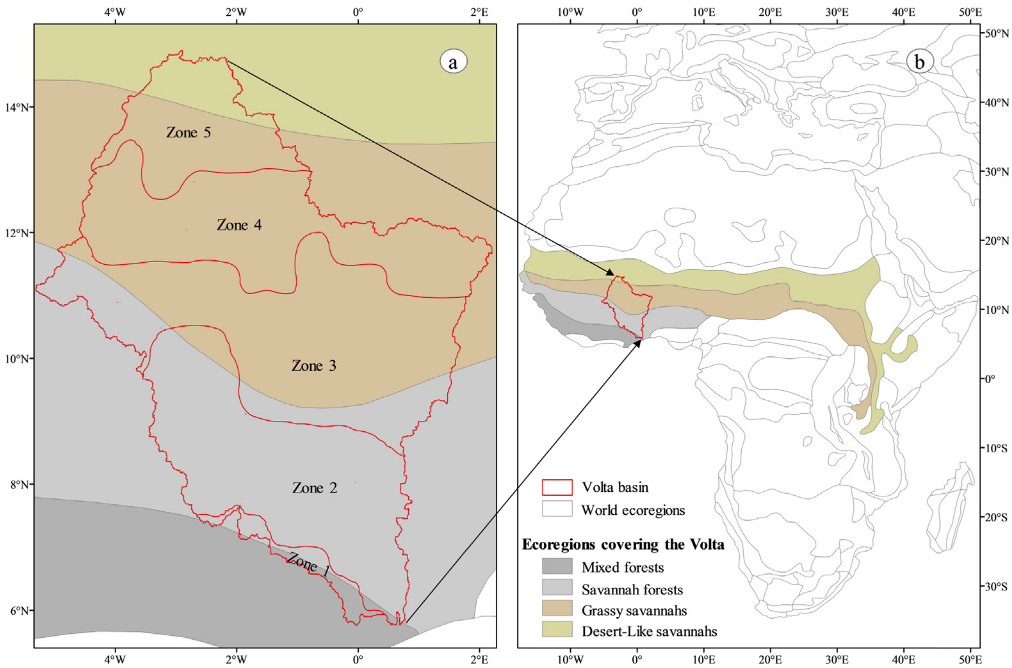

Fig. 7. Comparing the delineated zones of the Volta basin (a) with the United Nation Environmental Program's World ecoregions (b). The ecoregions encompassing the Volta basing are: (i) "Mixed forests" described as mixed forests with short dry season; (ii) "Savannah forests" described as humid tall-grass savannahs and savannah forests; (iii) "Grassy savannahs" described as moderately humid grassy savannahs; (iv) "Desert-like savannahs" described as desert-like savannahs, open woodland, and shrubs.

humid grassy savannahs; (iv) "Desert-like savannahs" described as desert-like savannahs, open woodland, and shrubs. The overlay of the Volta zones on the UNEP's ecoregions shows relevant similitudes. Indeed, the mixed forests ecoregion is likely represented by the zone 1 of the Volta, the savannah forests ecoregion is essentially represented by the zone 2 of the Volta. However, the grassy savannahs ecoregion overlaps the zones 3, 4, and a part of zone 5 of the Volta basin. The desert-like savannah is represented by the upper side of zone 5. From a general view, the new Volta zones are partially concordant with the world ecoregions, but the disparities noted are understandable given the fact that these ecoregions are the product of a macroscale assessment.

5. Discussion

Since Shannon (1948) introduced the entropy theory, the concept has been widely used in multiple fields of science and engineering (Singh, 2013). In environmental studies, the pertinence of the entropy theory was consistently demonstrated with practical implications (Singh, 2013,2011). The present study substantiates the usefulness of the entropy theory by describing its capacity to capture meaningful biophysical signals across a given spatial domain. Unlike ordinary spatial regionalization techniques, the method developed in this study intends to recover simultaneously spatial and temporal patterns within a multivariate framework. In reality, it is difficult to conceive a multivariate spatial regionalization using a one-dimension clustering method (Fischer et al., 2002). Yet, this does not mean that one-dimension clustering techniques are irrelevant. The zoning approach of this paper may be regarded as a potential tool for performing multivariate spatial regionalization tasks. The technique is developed for applications at the watershed scale. Precisely in this study, the Volta watershed was targeted then assessed using the proposed zoning technique. The zoning approach values the watershed as a functional unit.

Regardless of their size or morphology, watersheds generally host natural ecosystems and human communities as well (Castro et al., 2016; Sala et al., 2000). In several locations of the globe, these ecosystems are continuously affected by anthropogenic actions which are widely recognized as a major threat to environmental sustainability (Pecl et al., 2017). In the Volta river basin, environmental concerns are multifaceted and diversely perceived by the populations across the watershed (Rodgers et al., 2006; van de Giesen et al., 2001). Indeed, the Volta watershed is shared by six different countries which communities do not necessarily face the same environmental challenges (Oguntunde et al., 2017; Rodgers et al., 2006). For instance, several communities living in the upstream and middle-stream regions of the watershed are known for using important quantities of agricultural chemicals. As a result, their agricultural practices are frequently identified as a factor causing water impairments in the downstream regions of the Volta river basin (Yidana and Yidana, 2010; Karikari and Ansa-Asare, 2006; Ntow, 2005). To properly address such problem inherent to the Volta basin, it would be beneficial to envisage a consensual management strategy at the watershed scale (Jägerskog et al., 2012; van de Giesen et al., 2001). From that point of view, the contents of this paper are timely and relevant for the Volta river basin. The zoning technique developed for the study, employed gridded time-series of precipitation, temperature, and LAI. Note, precipitation, and temperature are two fundamental variables, generally used for characterizing climate variability (Almazroui et al., 2015). The variability of these two climate factors shows a northward gradient which is also reflected by the LAI distribution across the watershed (Fig. 3). Yet this common northward gradient is proper to the Volta watershed and was previously highlighted by Rodgers et al. (2006). However, meaningful disparities are brought out when one compares the extent of LAI distribution with the spatial variability of the climate factors.

In general, the variation of LAI in time and space is influenced by the climate but it also reflects the magnitude of anthropogenic factors such as agriculture, deforestation, urbanization (Huete et al., 2010; Garrigues et al., 2008; Xavier and Vettorazzi, 2004). Interestingly, the zoning framework of the Volta basin has captured critical patterns of LAI. Hence, Table 2 describes the main features of LAI in each of the five zones delineated. For instance, a gradual decrease of the LAI's range is observed from zone 1 to zone 5. In practice, knowing the range of LAI in a given location, one can portend approximatively the architecture of the dominant vegetation covers (Jonckheere et al., 2004). In fact, LAI defines the structural property of vegetation canopy and its value is interpreted as the number of equivalent foliage layers by unit area (Garrigues et al., 2008). Literally, high values of LAI are associated to multi-layer vegetation covers (Leblanc and Fournier, 2017). From this standpoint, the vegetation cover in zone 1 can be labeled as multi-stratified compared to the other zones. In reality, each of the five delineated zones can be linked to specific vegetation formations in the West Africa region. Hence, zone 1 is part of the deciduous forest region of Ghana, zone 2 is transitional, zone 3 corresponds mainly to a savannah (i.e. the Guinea savannah), while zone 4 and 5 are dominated by sparse Sahelian vegetation (Anyamba and Tucker, 2005; Windmeijer and Andriesse, 1993). This contrast of land-cover between the zones aligns with the general precipitation pattern of the Sahelian region (Nicholson, 2013). The land-cover contrast may also explain the paradigm portrayed by Fig. 6 which showed different responses of LAI to precipitation among the zones. Depending on the types, land-covers are generally impacted differently by the cross-seasonality which results from the annual succession of dry and moist seasons (e.g. Fig. 5: dry season = November to May, moist season = June to September). For instance, savannahs are likely to exhibit a non-cumulative cross-seasonal effect while forests are likely to respond differently by exhibiting cumulative cross-seasonal effects (Sohoulande Djebou et al., 2015).

The comparison of the new Volta zones with the UNEP's world ecoregions shows critical resemblances as the zone 1 and 2 were essentially assimilated to the mixed forests and the savannah forests ecoregions respectively (Fig. 7). However, the UNEP's ecoregions are macro-sales subdivisions which are likely to merge disparities at finer scales. Thus, one can understand that zones 3, 4 and 5 are merged in the same macro-ecoregion (i.e. grassy savannahs). Note, the data used to develop the Volta basin zones are more recent and are likely to provide critical insights into the ongoing environmental changes across the watershed (Ndehedehe et al., 2017; Andam-Akorful et al., 2015). Indeed, a recent study reports an overall freshwater depletion across the Volta river basin (Ndehedehe et al., 2017), but a drought assessment highlights contrasting drought zones across the watershed (Ndehedehe et al., 2016). These results sustain the need for a unified zoning framework for the Volta watershed.

Overall, the land-cover and climatic contrasts described across the Volta watershed, seem to substantiate the consistency of the regionalization technique developed in this study. Precisely, the outcomes of the study, demonstrate the capacity of the entropy theory and the k-means clustering to detect meaningful environmental patterns at the watershed scale. In view of the paradigm portrayed by Fig. 6, it is normal to also assert the capacity of the zoning technique to pinpoint biophysical interactions at the watershed scale. These results align very much with previous studies which reported the inclusion of entropy theory in spatial regionalization (Tongal and Sivakumar, 2017). However, the identification of the optimal number of zones was facilitated by using the cubic clustering criterion during the k-means clustering stage (Ketchen and Shook, 1996; Milligan and Cooper, 1985). Ultimately, the zoning method developed is overall consistent with the primary objective of the study and it could, therefore, be considered as an environmental management tool.

6. Conclusion

The spatial regionalization technique employed in this study helped to identify five zones with meaningful environmental specificities across the Volta watershed. The zones were comparatively analyzed, then significant statistical differences in precipitation, temperature, and LAI patterns were reported. The overall performance of the zoning technique is satisfactory as it uncovered signals of spatial heterogeneity between the delineated zones. Furthermore, relevant similitudes were highlighted between the Volta zones and the world ecoregions encompassing the Volta basin. However, the Volta zoning framework seems to bring out local details which were apparently merged within the macro-scale world ecoregions. As a result, the developed zoning framework may serve as a technical tool for enhancing ecosystem management in the Volta river basin. Indeed, several studies emphasized the imperative of enhancing the integrated management strategies in the Volta river basin (Rodgers et al., 2006; van de Giesen et al., 2001). Aligning with such imperative, it is important to establish a unified zoning system which may overcome the challenges inherent to the transnational status of the Volta watershed (Oguntunde et al., 2017; Jägerskog et al., 2012). The zoning system reported in this study offers such a unified platform. Therefore, the delineated zones may be incorporated in the environmental management agenda of the Volta watershed.

Beyond the reported case study, the value of this paper lies also in its methodological contribution. Indeed, the theoretical zoning approach is explicitly presented and described in the methodology section of the paper. Thus, the author encourages an extensive use of this regionalization technique for watershed studies. However, the use of the approach may be subjected to modifications depending on data availability and the finality targeted by the user.

Disclaimer

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

Abell, R., Thieme, M.L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S.C., Bussing, W., Stiassny, M.L., 2008. Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. AIBS Bull. 58 (5), 403–414.
Almazroui, M., Dambul, R., Islam, M.N., Jones, P.D., 2015. Principal components-based regionalization of the Saudi Arabian climate. Int. J. Climatol. 35 (9), 2555–2573.
Andam-Akorful, S.A., Ferreira, V.G., Awange, J.L., Forootan, E., He, X.F., 2015. Multi-model and multi-sensor estimations of evapotranspiration over the Volta Basin, West Africa. Int. J. Climatol. 35 (10), 3132–3145.
Antwi, E.K., Boakye-Danquah, J., Asabere, S.B., Yiran, G.A., Loh, S.K., Gyekye, K.A., Abagale, F.K., Asubonteng, K.O., Attua, E.M., Owusu, A.B., 2014. Land Use and Landscape Structural Changes in the Ecoregions of Ghana.
Anyamba, A., Tucker, C.J., 2005. Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. J. Arid Environ. 63 (3), 596–614.
Bailey, R.G., 1989. Explanatory supplement to ecoregions map of the continents. Environ. Conserv. 16 (4), 307–309.
Bailey, R.G., Hogg, H.C., 1986. A world ecoregions map for resource reporting. Environ. Conserv. 13 (3), 195–202.
Bharati, L., Rodgers, C., Erdenberger, T., Plotnikova, M., Shumilov, S., Vlek, P., Martin, N., 2008. Integration of economic and hydrologic models: exploring conjunctive irrigation water use strategies in the Volta Basin. Agric. Water Manag. 95 (8), 925–936.
Brou, Y.T., Akindès, F., Bigot, S., 2005. La variabilité climatique en Côte d'Ivoire: entre perceptions sociales et réponses agricoles. Cah. Agric. 14 (6), 533–540.
Castro, A.J., Vaughn, C.C., Julian, J.P., García-Llorente, M., 2016. Social demand for ecosystem services and implications for watershed management. JAWRA J. Am. Water Resour. Assoc. 52 (1), 209–221.
Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A.S., McLoone, S., 2013. Principal component analysis on spatial data: an overview. Ann. Assoc. Am. Geogr. 103 (1), 106–128.
Fischer, G., van Velthuizen, H.T., Shah, M.M., Nachtergaele, F.O., 2002. Global Agro-Ecological Assessment for Agriculture in the 21st Century: Methodology and Results. IIASA Research Report. IIASA, Laxenburg, Austria, pp. 1–119.
Garrigues, S., Lacaze, R., Baret, F.J.T.M., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., 2008. Validation and intercomparison of global Leaf Area Index products derived from remote sensing data. J. Geophys. Res. Biogeosci. 113 (G2).
Harris, I.C., Jones, P.D., 2017. CRU TS4.00: Climatic Research Unit (CRU) Time-Series (TS) version 4.00 of high resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2015). Centre for Environmental Data Analysis, 25 May 2017 ⟨https://www.doi.org/10.5072/edf8febfdaad48abb2cbaf7d7e846a86⟩.
Hornbeck, J.W., Swank, W.T., 1992. Watershed ecosystem analysis as a basis for multiple-use management of eastern forests. Ecol. Appl. 2 (3), 238–247.
Huete, A., Didan, K., van Leeuwen, W., Miura, T., Glenn, E., 2010. MODIS vegetation indices. In: Ramachandran, B., Justice, C.O., Abrams, M.J. (Eds.), Land Remote Sensing and Global Environmental Change. Springer, New York, pp. 579–602.
Jägerskog, A., Cascão, A., Hårsmar, M., Kim, K., 2012. Land Acquisitions: How Will They Impact Transboundary Waters. Stockholm International Water Institute, Stockholm, pp. 1–32.
Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F., 2004. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. Agric. For. Meteorol. 121 (1), 19–35.
Karikari, A.Y., Ansa-Asare, O.D., 2006. Physico-chemical and microbial water quality assessment of Densu River of Ghana. West Afr. J. Appl. Ecol. 10 (1).
Ketchen Jr, D.J., Shook, C.L., 1996. The application of cluster analysis in strategic management research: an analysis and critique. Strateg. Manag. J. 441–458.
Knyazikhin, Y., Glassy, J., Privette, J.L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J.T., Votava, P., Myneni, R.B., Nemani, R.R., 1999. MODIS Leaf Area index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document 20771. Theoretical Basis Document, NASA Goddard Space Flight Center, Greenbelt, MD, pp. 126.
Leblanc, S.G., Fournier, R.A., 2017. Measurement of forest structure with hemispherical photography. In: Fournier, R.A., Hall, R.J. (Eds.), Hemispherical Photography in Forest Science: Theory, Methods, Applications. Springer, Netherlands, pp. 53–83.
Milligan, G.W., Cooper, M.C., 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50 (2), 159–179.
Ndehedehe, C.E., Awange, J.L., Corner, R.J., Kuhn, M., Okwuashi, O., 2016. On the potentials of multiple climate variables in assessing the spatio-temporal characteristics of hydrological droughts over the Volta Basin. Sci. Total Environ. 557, 819–837.
Ndehedehe, C.E., Awange, J.L., Kuhn, M., Agutu, N.O., Fukuda, Y., 2017. Analysis of hydrological variability over the Volta river basin using in-situ data and satellite observations. J. Hydrol. Reg. Stud. 12, 88–110.
Nicholson, S.E., 2013. The West African Sahel: a review of recent studies on the rainfall regime and its inter annual variability. ISRN Meteorol. 2013.
Ntow, W.J., 2005. Pesticide residues in Volta Lake Ghana. Lakes Reserv. Res. Manag. 10 (4), 243–248.
Oguntunde, P.G., Abiodun, B.J., Lischeid, G., 2017. Impacts of climate change on hydro-meteorological drought over the Volta Basin, West Africa. Glob. Planet. Change 155, 121–132.
Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.C., Clark, T.D., Colwell, R.K., Danielsen, F., Evengård, B., Falconi, L., 2017. Biodiversity redistribution under climate change: impacts on ecosystems and human well-being. Science 355 (6332), eaai9214.
Rodgers, C., van de Giesen, N., Laube, W., Vlek, P.L., Youkhana, E., 2006. The GLOWA Volta Project: a framework for water resources decision-making and scientific capacity building in a transnational West African basin. In: Craswell, E., Bonnell, M., Bossio, D., Demuth, S., van de Giesen, N. (Eds.), Integrated Assessment of Water Resources and Global Change. Springer, Netherlands, pp. 295–313.
Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., 2000. Global biodiversity scenarios for the year 2100. Science 287 (5459), 1770–1774.
Santos, M., Fragoso, M., Santos, J.A., 2017. Regionalization and susceptibility assessment to daily precipitation extremes in mainland Portugal. Appl. Geogr. 86, 128–138.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (379–423), 623–656.
Singh, V.P., 2013. Entropy Theory and its Application in Environmental and Water Engineering. Publisher, Wiley, pp. 666.
Singh, V.P., 2011. Hydrologic synthesis using entropy theory: review. J. Hydrol. Eng. 16 (5), 421–433.
Sohoulande Djebou, D.C., Singh, V.P., Frauenfeld, O.W., 2015. Vegetation response to precipitation across the aridity gradient of the southwestern United states. J. Arid Environ. 115, 35–43.
Sturges, H.A., 1926. The choice of a class interval. J. Am. Stat. Assoc. 21 (153), 65–66.
Tongal, H., Sivakumar, B., 2017. Cross-entropy clustering framework for catchment classification. J. Hydrol. 552, 433–446.
van de Giesen, N., Andreini, M., van Edig, A., Vlek, P., 2001. Competition for Water Resources of the Volta Basin. IAHS Publication, pp. 199–206.
Windmeijer, P.N., Andriesse, W., 1993. Inland Valleys in West Africa: an Agro-ecological Characterization of Rice-growing Environments. 52 ILRI.
Xavier, A.C., Vettorazzi, C.A., 2004. Mapping leaf area index through spectral vegetation indices in a subtropical watershed. Int. J. Remote Sens. 25 (9), 1661–1672.
Yidana, S.M., Yidana, A., 2010. Assessing water quality using water quality index and multivariate analysis. Environ. Earth Sci. 59 (7), 1461–1473.
Yoon, K.A., Kwon, O.S., Bae, D.H., 2007. An approach to outlier detection of software measurement data using the k-means clustering method. In: Proceedings of the First International Symposium on Empirical Software Engineering and Measurement, ESEM 2007. IEEE, September, pp. 443–445.
Yue, S., Pilon, P., Cavadias, G., 2002. Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrol. 259 (1), 254–271.
Zhang, Y., Moges, S., Block, P., 2016. Optimal cluster analysis for objective regionalization of seasonal precipitation in regions of high spatial–temporal variability: application to Western Ethiopia. J. Clim. 29 (10), 3697–3717.
Zhang, Y.Q., Chiew, F.H.S., Zhang, L., Leuning, R., Cleugh, H.A., 2008. Estimating catchment evaporation and runoff using MODIS leaf area index and the Penman-Monteith equation. Water Resour. Res. 44 (10).
Zhu, Z., Piao, S., Myneni, R.B., Huang, M., Zeng, Z., Canadell, J.G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., 2016. Greening of the Earth and its drivers. Nat. Clim. Change 6 (8), 791–795.