Characterization of Soil Degradation from the Cameroonians Shores of Lake Chad Combining Spectral Indexes and Statistics Analysis

Sébastien Gadal, Paul Gérard Gbetkom, Alfred Homère Ngandam Mfondoum

SN Computer Science (2023) 4:237
https://doi.org/10.1007/s42979-022-01651-7

ORIGINAL RESEARCH

Characterization of Soil Degradation from the Cameroonians Shores of Lake Chad Combining Spectral Indexes and Statistics Analysis

Sébastien Gadal[1,2] · Paul Gérard Gbetkom[1] · Alfred Homère Ngandam Mfondoum[3]

Received: 4 October 2021 / Accepted: 21 December 2022

Abstract

This article aims to propose a model on the soil degradation risk along the Cameroonian shores of Lake Chad based on the statistical analysis of spectral indexes of Sentinel 2A satellite images. A total of four vegetation indexes such as the Greenness Index and Disease water stress index and nine soil indexes such as moisture, brightness, or organic matter content are computed and combined to characterize vegetation cover and bare soil state, respectively. All these indexes are aggregated to produce one image (independent variable) and then regressed by individual indexes (dependent variable) to retrieve correlation and determination coefficients. Principal Component Analysis and factorial analysis are applied to all spectral indexes to summarize information, obtain factorial coordinates, and detect positive/negative correlation. The first factor contains soil information, whereas the second factor focuses on vegetation information. The final equation of the model is obtained by weighting each index with both its coefficient of determination and factorial coordinates. This result generated figures' cartography of five classes of soils potentially exposed to the risk of soil degradation. Five levels of exposition risk are obtained from the "Lower" level to the "Higher": the "Lower" and "Moderate to low" levels occupy, respectively, 25,214.35 hectares and 130,717.19 hectares; the "Moderate" level spreads 137,404.34 hectares; the "High to moderate" and "Higher" levels correspond, respectively, to 152,371.91 hectares and 29,175.73 hectares.

Keywords Soil degradation · Vegetation indexes · Soil indexes · Statistical analysis · Lake Chad · Sentinel 2A

Introduction

The state of soil is an important parameter in the monitoring of the land dynamic and exploitation, for sustainable use [12, 27]. Its degradation, which reduces the exploitation of natural resources in general and restricts the productivity of agricultural soils, causes significant socio-economic impacts. In the far-northern part of Cameroon, the shores of Lake Chad are in the most exposed zone to the soil degradation risks due to environmental conditions, more severe climatic conditions, and modes of use and exploitation of natural resources [National Action Plan to Combat Desertification (PAN/LCD) [46]]. It is an area marked by degradation and decline of soil fertility, unsuitable cultivation practices, a high extension of barren land, erosion, runoff, and decrease of fallows, overgrazing, and pesticide pollution (Elias [23, 59].

The great spatial and temporal variability of the rainfall combined with the rain aggressiveness constitutes major risks related to the rainfall and accelerates the soil degradation process in this zone (PAN/LCD, 2006).

This article is part of the topical collection "Geographical Information Systems Theory, Applications and Management" guest edited by Lemonia Ragia, Cédric Grueau and Robert Laurini.

✉ Sébastien Gadal
sebastien.gadal@univ-amu.fr

Paul Gérard Gbetkom
paul.gerard.gbetkom@legos.obs-mip.fr

Alfred Homère Ngandam Mfondoum
stats.n.maps.expertise@gmail.com

[1] Aix-Marseille University, CNRS, University Nice Sophia Antipolis, Avignon University, ESPACE UMR 7300, 13545 Aix-en-Provence, France

[2] North-Eastern Federal University, 670000 Yakutsk, Republic of Sakha-Yakutia, Russia

[3] StatsN'Maps, Private Consulting Firm, 19002 Dallas Parkway, Suite 1536, Dallas, TX 72587, USA

Published online: 02 March 2023

Rainfalls as violent localized showers, and strike bared soils, prepared for sowing and lowly protected or cleaned from their vegetation [55]. This is figured out by the presence of vast expanses of bare soils, most of which are very sensitive to water and wind erosion, accentuated by the dwindling vegetation cover. Slopes are low in this environment, and the level of soil drainage is very varied. It is moreover based on the level of draining that [55] distinguished the well-drained lands (terroir of Makari), the poorly drained, lands with waterlogging (terroir of Bodo-Kouda), and the poorly drained lands with waterlogging and fluvial (terroir of Lake Chad). It is a periodically flooded area, where the main activities are fishing, livestock, agriculture, and trade, shared by a large and varied population coming from at least four neighbouring countries (Cameroon, Chad, Niger, and Nigeria), with consequently numerous conflicts.

Several methods are used to quantify and map soil degradation at different spatial and temporal scales. Universal Soil Loss Equation [61] or its modified version [52] is used to predict soil erosion. This model depends on the slope, the rainfall, the soil typology, topography, crop rotation, and the soil conservation practice. Further, Ali et al. [2] have used sedimentology and magnetic measurements to identify sediment source areas, assess spatial variations in sediment levels, and classify these zones depending on their degree of spatial reworking. Another method was proposed by Daniel et al., 2018 to map the soil degradation, by collecting field samples and performing an unsupervised Iterative Self-Organizing Data Analysis Technique (ISODATA) classification on the combination of Sentinel-2 data image and airborne orthoimages. Richard and Huete [18] proposed a monitoring approach for arid land surface status, based on satellite images. The United Nations Convention to Combat Desertification quantifies soil organic carbon and extract indicators as soil productivity and land cover using MODIS NDVI data, to map the proportion of land degraded over the world (support by Conservation International, Lund Université, NASA, and Global Environment Facility).

All the above-described methods include several ancillary data and field samples of the study area and need to consider the topography of the field. However, the ancillary data are not available in the total to the extent of our study area and the distinct types of soil topology and topography are not easy to distinguish due to the spatial resolution of the image used. Therefore, we need to develop a new remote-sensing approach only based on the soil and the vegetation spectral indexes which can allow identifying areas exposed to the risk of soil degradation.

Indeed, remote sensing enables collecting and integrating data for a continuous and repeated observation of the phenomenon on large surfaces [22]. The reflectance of some objects such as soil and vegetation is a good indicator of changes in the environment [21] and can be used to calculate spectral indexes useful for the study of soil degradation. Previous models have been developed on the topic. It is the case of Ngandam et al. [40] who use the linear and the multiple regressions, and the principal component analyses to assess the status of soil degradation in Far-North Cameroon. Following this last work, the statistical methods will be supplemented in this work by other statistical processes, such as factor analysis, to highlight the level of correlation between the selected indexes. Thus, the indexes, such as the Normalized difference vegetation index (NDVI), Modified Soil Adjusted Vegetation Index (MSAVI2), Normalized Difference Greenness Index (NDGI), and Disease water stress index (DSWI), are used in this study to quantify vegetation cover and provide information, respectively, on chlorophyll activity, the density of vegetation cover, vegetation greenery, and plant water stress. On the other hand, soil characteristics are highlighted through spectral indexes. Those used for that purpose in this study are moisture stress index (MSI), texture index (TI); colour index (IC); brightness index (BI); cuirass index (CI); topsoil surface particles index (GSI); crusting index (CI); redness index (RI); and salinity index (NDSI).

The mapping of soil degradation from indexes is sometimes limited to a simple combination of the index in the form of a band-coloured composition [35] or to an approach that associates spectral indexes with different classification methods [14]. Atman et al. [1] thus use spectral indices for the assessment of desertification in arid oasis. On the other hand, cross-indexes and models soil degradation by weighting indexes and neo-bands using the coefficient of determination resulting from the linear regression between each index and the weighted sum image [40]. In their approach, [44] cross-spectral indexes to land-cover maps, but index maps are reclassified according to the level of severity of land degradation and associated with land-use and land-cover map. Therefore, this paper explores another modelling approach to assess soil degradation. Specifically, in three steps, it highlights soil properties through spectral indexes. After that, it proceeds to a statistical analysis of the index contents to withdraw their correlation trends. Finally, the two steps above propose an overall model to predict soil degradation risk.

Methodology

The Study Area: Cameroonian Part of Lake Chad and Hinterland

The study area is located in the Far north administrative region of Cameroon and shares borders with the Republic of

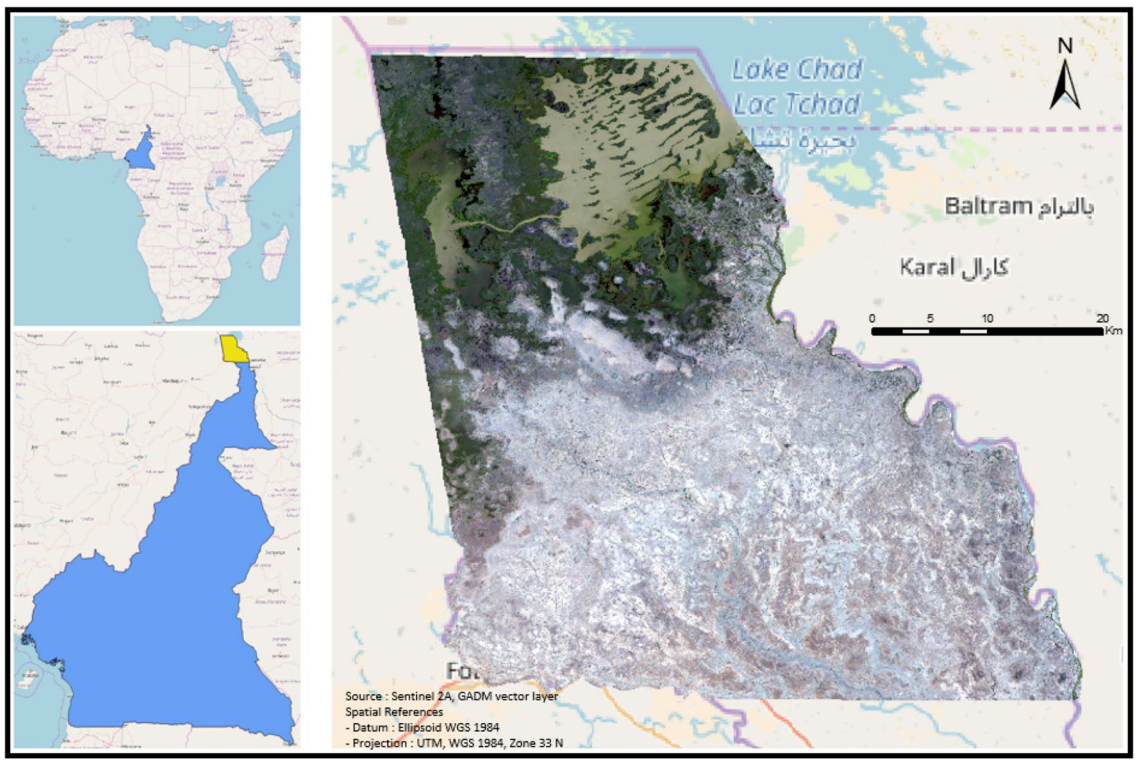

Fig. 1 Localisation of the study area

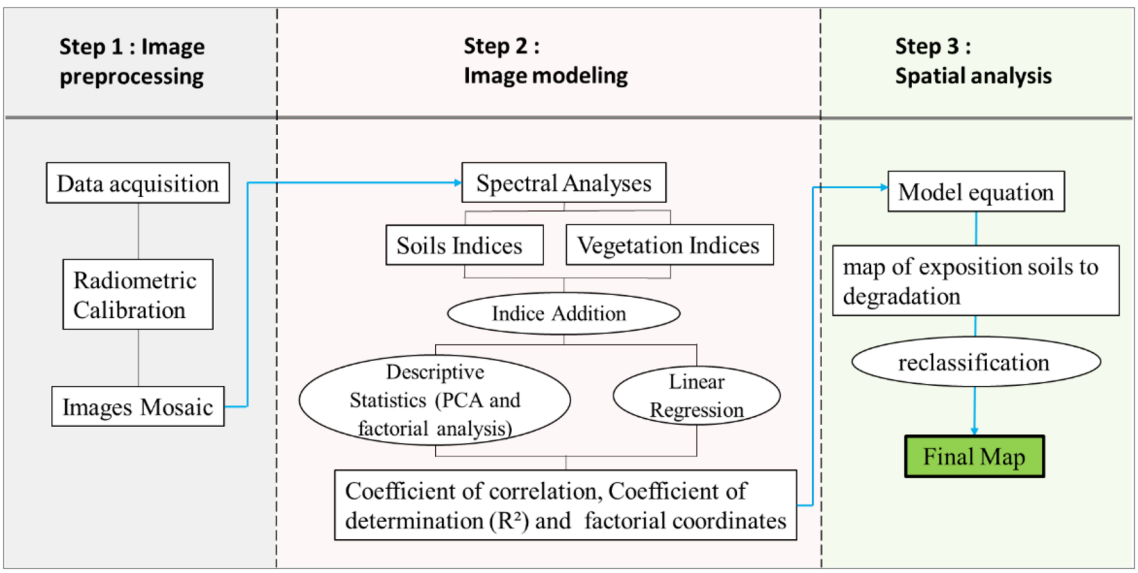

Fig. 2 Flowchart of the methodology

Table 1 Characteristic of Sentinel 2A images

Spatial resolution	Band number	Band name	Wavelength (µm)
10 m	B02	Blue	490
	B03	Green	560
	B04	Red	665
	B08	NIR	842
20 m	B05	Red Edge 1	705
	B06	Red Edge 2	740
	B07	Red Edge 3	783
	B08A	Red Edge 4	865
	B11	SWIR 1	1610
	B12	SWIR2	2190
60 m	B01	Aerosol	443
	B09	Water vapour	940
	B10	Cirrus	1375

Chad to the north and east, the Federal Republic of Nigeria to the west, and the rest of the country to the south (Fig. 1). It is located between latitude 12°N to 13°N and meridian 14°E to 15°E. It is a semi-arid region with a Sudano-Sahelian climate, characterized by a rainy season from June to October and a dry season that runs from November to May. The annual rainfall totals around 400 mm, the temperature range is 7.7 °C, and the average monthly temperature is 28 °C.

An Approach Based on Sentinel-2 Images

Two Sentinel 2A satellite images acquired on April 29, 2017 were used (Fig. 2). They have 13 bands, but only six of them were staked, i.e., Bands 2, 3, 4, 8, 11, and 12, as detailed in Table 1.

Soil Risk of Degradation Model Design

Soil and Vegetation Indexes Modelling

a. The Vegetation Indexes The use of vegetation indexes has several objectives, such as the estimation of the green vegetable mass, the forecast of harvests, the description of the phonological state of the soil cover, the inventory of crops by segmentation of indexes, and the evolution of vegetation cover at the continental scale [9]. For example, the ability of the NDVI to detect the presence, density, and condition of vegetation was successfully used by [15], to observe a regreening of the Sahel between 1982 and 2003, due to the spatial increase in vegetation cover. The following indexes were therefore used in this study to analyze the chlorophyll activity, the density of vegetation cover, the vegetation greenery, and plant water stress (Table 2).

The visual comparison of the vegetation index's efficiency to discriminate and quantify canopy density shows a more accurate representation using MSAVI2. Unlike the NDVI, the MSAVI2 offers a sensitive distinction between bare soils and green areas in less-vegetated regions. Also, this index attributes low values to aquatic spaces in contrast to the DSWI and NDGI indexes. These observations are consistent with the results of previous works that showed the potential of MSAVI2 to map the state of the vegetation cover in arid environments [19, 40]. The four indexes distinguish vegetated areas from bare soils. However, the use of soil indexes in addition to vegetation indexes is essential to characterize the bare spaces. Therefore, nine soil indexes are computed and combined (Fig. 3).

b. The Soil Indexes Escadafal and Huete [17] use the soil colour index to distinguish surface materials from soils according to the saturation of their colour. Chikhaoui et al.

Table 2 Characteristics of the vegetation indexes

Indexes	Algorithm	Goal	References
For the chlorophyll activity: NDVI	$NDVI = (NIR - R) / (NIR + R)$ [54]	Used to evaluate the chlorophyll activity of plants and also for the monitoring of the state of the vegetation cover	[37], (E. [45, 58], (Farooq [19])
For the density of vegetation cover: MSAVI2	$MSAVI\,2 = \frac{2NIR+1-\sqrt{(2NIR+1)^2-8(NIR-R)}}{2}$ [49]	Description of the vegetation density and reduces the effects of soil, in particular when the canopy is sparse especially in arid and semi-arid environments	[40, 49], (Farooq [19])
For the characterization of the plant water stress: DSWI	$DSWI = (NIR + G)/(SWIR + R)$[4]	Used to describe the variation of the water content of foliage	[48], X. [4, 32]
For the recognition of the vegetation greenery: NDGI	$NDGI = (G - R)/(G + R)$[10]	Used to estimate the biomass of vegetation and measure the hydric potential of the leaves at the level of the canopy	[20, 53], (H. [31, 50], (Sun, Li, and Li [57])

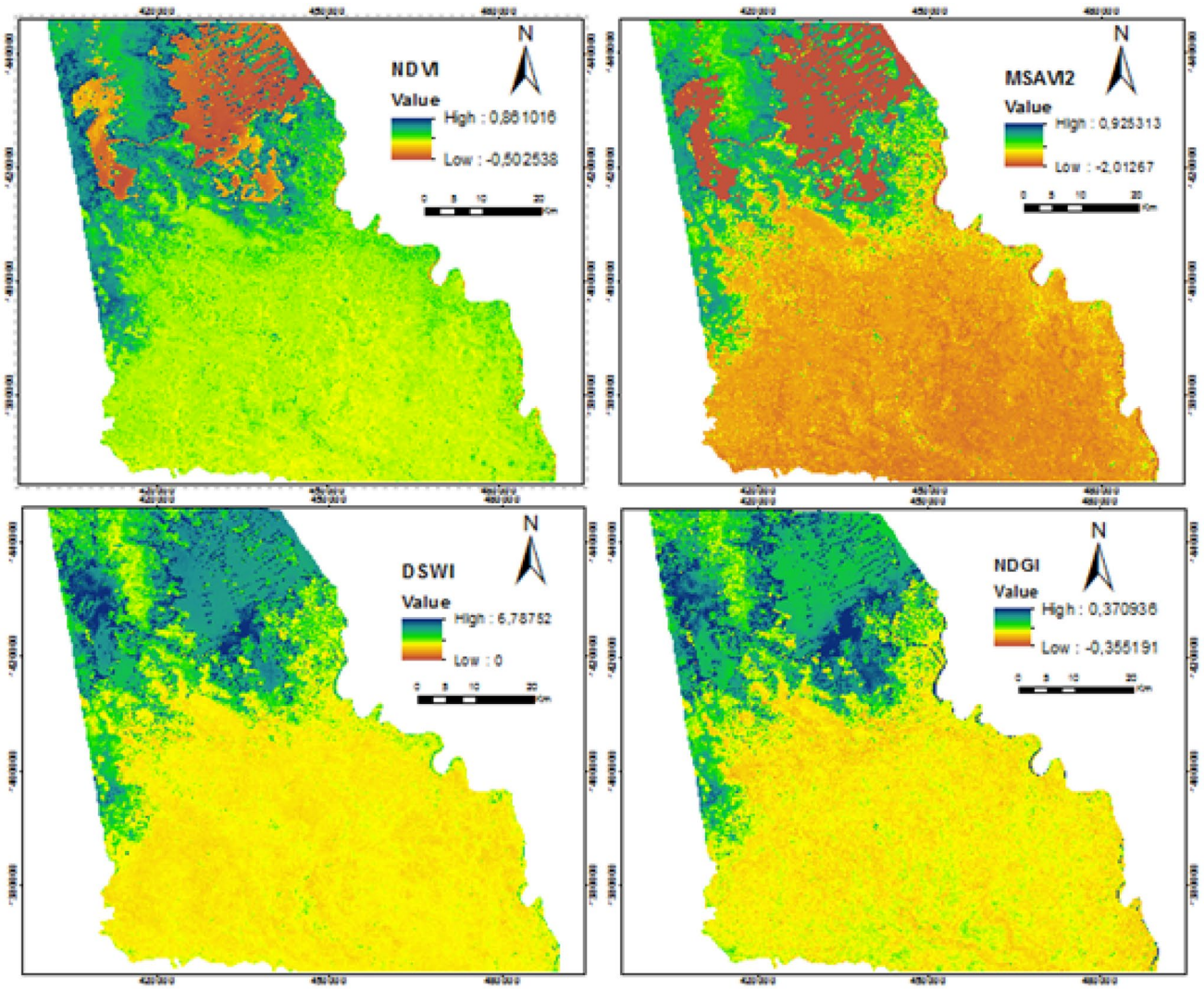

Fig. 3 The vegetation indexes

[13] characterize the state of land degradation in Morocco through the land degradation index (LDI).

The following indexes were therefore used in this study to highlight the mineralogical composition of soils, and to assess the organic matter content of soils, and the physical state of soils in terms of moisture and compactness. Moreover, parameters, such as colour, brightness, texture, and moisture, characterize the absorption properties of the soil constituents and are important for mapping soil conditions, particularly in arid environments (Table 3).

The indexes characterize the soils using the reflectance curves and the spectral properties of the soil constituents (MSI, BI, crust Index, TI, cuirass Index, RI, colour Index, GSI, and NDSI) (Fig. 4). Cuirass and crust indexes show that compact soils are mostly present in the southern part of the study area where soils are completely bare. The low values of the colour index coincide with the high values of the redness index and correspond to the densely vegetated Lake Chad littoral spaces, which are therefore rich in organic matter. On the other hand, spaces with a low redness index have high values of brightness, MSI, and NDSI, which indicates low soil moisture and a prominent level of drought and soil salinity. Furthermore, in the southern part of the study area, where the levels of cuirass and crust are already high, the soil texture is also dominated by the presence of coarse particles considering the results of the texture indexes and GSI.

Table 3 Characteristics of the soil indexes

Indexes	Algorithm	Goal	References
The moisture stress: MSI	MSI = SWIR1/NIR(Yongnian et al. [64])	Used to evaluate the spatial extend of less soil moisture, due to the higher level of evapotranspiration	[16, 60]
The texture analysis: TI	TI = (SWIR1 − SWIR2)/(SWIR1 + SWIR2) (Madeira Netto [33])	The texture index is calculated to evaluate the content or percentage of sand, silt, and clay in soil composition, and appreciate the level of the mineral alteration of rock	(Madeira Netto [33]); [26, 42]
The soil colour: CI	CI = (R − V)/ (R + V)[17]	This index is used to extract information concerning the organic matter content and mineralogical composition of the soil	(Soufiane Maimouni and Bannari [35])
The soil brightness: BI	BI = $\sqrt{PIR^2 + R^2}$[29]	The role of the brightness index is to identify the reflectance of soil and to highlight the vegetal cover of bare areas	[8], (Soufiane [35]
The soil cuirass: CI	CI = 3 ∗ G − R − 100[47]	It aims to dissociate vegetated coverings from mineralized surfaces	[41], Stéphane et al. [56])
The topsoil grain size: GSI	GSI = (R − B)/(R + V + B)[62]	GSI or topsoil grain size index is an index appropriated to characterize the texture of the soil surface depending on the soil reflectance curve	(Jieying Xiao, Shen, and Ryutaro [63]); [40]
The soil crusting: CI	CI = (R − B)/(R + B)[28]	It is used to detect and map from satellite imagery different lithological morphological units. It is also able to reveal poor infiltration, and reduced air exchange between the soil and the atmosphere	[28]
The soil redness: RI	RI = $R^2/B * G^3$[38]	Used as one of the indicators to evaluate the mineralogy of soils, including the iron content	[17, 36, 51]
The soil salinity: NDSI	NDSI = (R − NIR)/(R + NIR)[30]	It is used to identify soils affected by salinity, and to show the spatial extent of salinity prevalent in our study area	[3, 5, 6, 11, 24], (Narmada, et al. [39])

Statistical Patterns

The model being developed also depends on the statistical information withdrawn from the indexes. This includes linear regression, factor analysis, and principal component analysis were calculated.

a. Linear Regression By adding all the indexes used, we obtain a new image that summarises all the information provided by each index. The image obtained will serve as the independent variable for the linear modelling between indexes. The purpose is to highlight the potential regressions between the image of the indexes used here as an explanatory variable, and each of the vegetation and soil indexes used as explanatory variables (Fig. 5).

Five indexes are negatively correlated to the synthetic image. These are the NDVI (−0.119), the DSWI (−0.747), NDGI (−0.769), TI (−0.606), and the redness index which has the most negative coefficient of correlation (−0.827). Moreover, the high values of the coefficient of determination of all the soil indexes except the NDSI show their influence on the synthetic image (Table 4). The other correlations are positive with values ranging from 0.119 for the NDSI to 0.953 for the brightness index.

The index most strongly determined by the synthesis image is the brightness index with an R^2 equal to 0.909. The other soil indexes have an R^2 with values that vary in

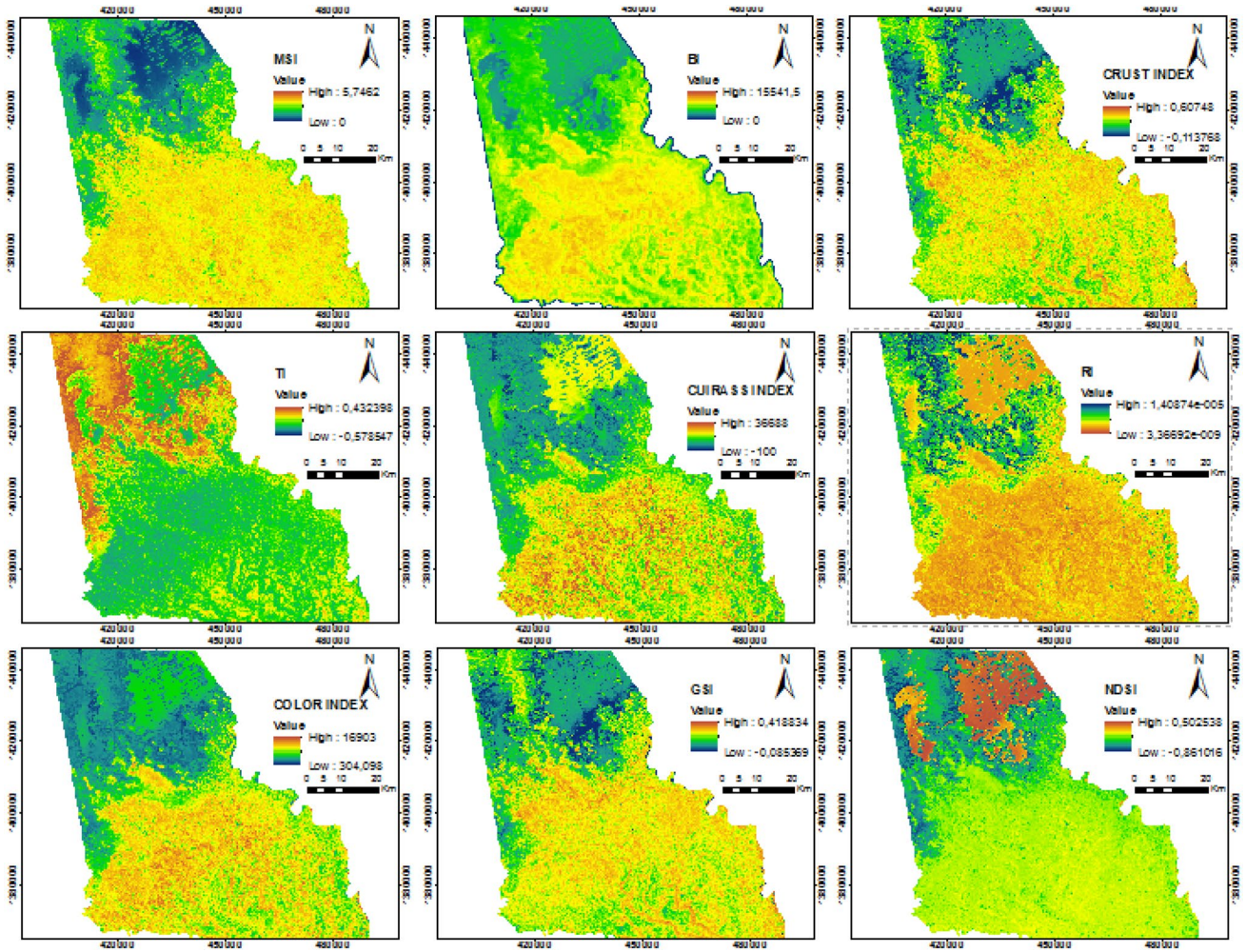

Fig. 4 The soil indexes

the interval [0.037–0.385]. For vegetation indexes, the R^2 values are contained between 0.007 and 0.592.

b. Descriptive Statistics For this step, we use factorial analysis which is a fundamental method of statistical analysis that does not have a particular structure [7, 43]. It is usually combined with the Principal Component Analysis (PCA) that is an extremely powerful data process for synthesizing information [25], to reduce dimensional space (two for example) to obtain the most relevant summary of the initial data. The output graphs are supported by characteristic numerical values, useful to ease the interpretation of the results. The graphs to be interpreted are the geographical representations and the tables which make it possible to see the connections and the oppositions between the studied characteristics, according to the factors used for illustration: the indexes of vegetations, the soil index, etc.

Factors "one" and "two" condense the most information and explain 85.62% of the common variability of the characteristics measured for the factor analysis and 87.54% for the PCA (Fig. 6). Moreover, for each method, factor one with more than 54% of the information is more important than factor two that contains a little more than 31%. For each factor, the best-revealed indexes are displayed in bold, and the opposition of the indexes is measured by the signs of the values (Table 5). For both methods, the first factor opposes DSWI, greenery, and redness indexes, with MSI, brightness, crusting, cuirass, colour, and GSI indexes. On the second factor, the NDVI, MSAVI2, and texture indexes are opposed to the NDSI index. The degrees of opposition and their disposition are illustrated by the graph of correlations between variables and factors (Fig. 6).

The symmetrical opposition of the first factor indexes shows that in the studied area when the greenery is high and the DSWI is also high, the brightness and the crusting of the soils decrease, the soils are wetter, darken, and the granulometry of topsoil is dominated by small

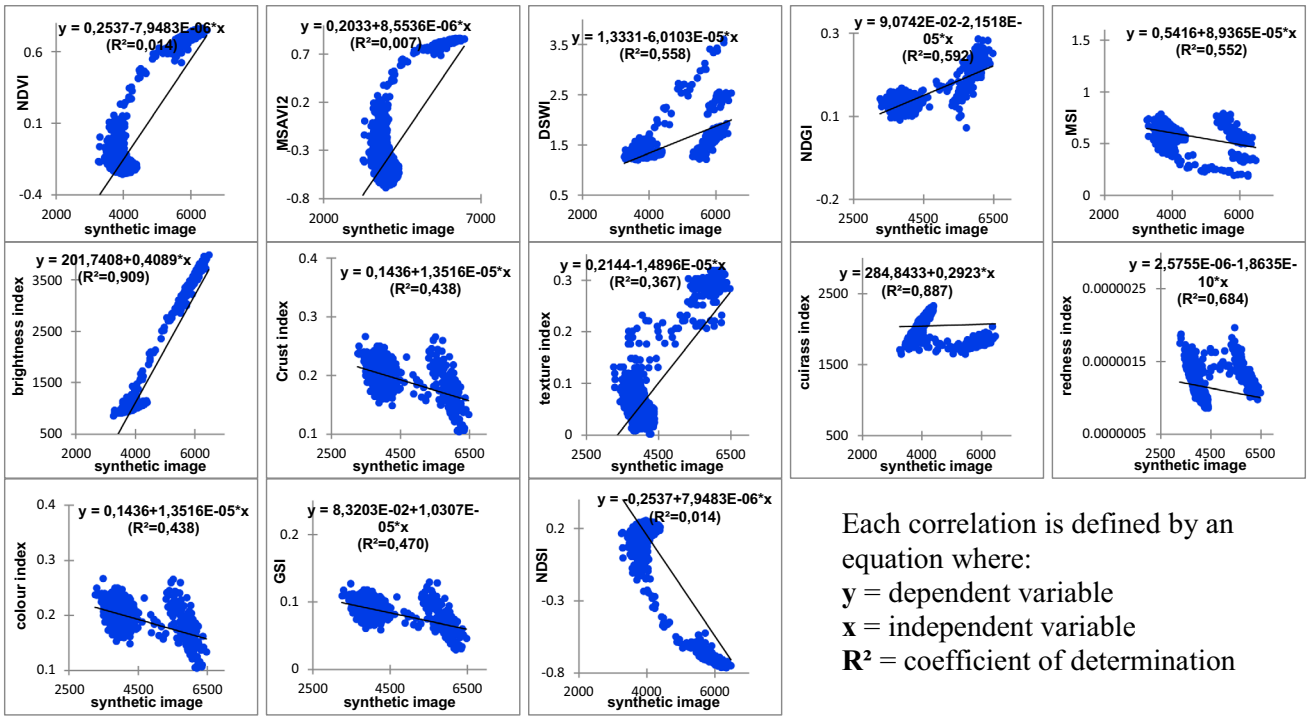

Fig. 5 Correlation between synthetic image and index's

Table 4 Statistics relations between synthetic image and indexes

Indexes	Correlation coefficient	Determination coefficient	*P* values	
			Threshold	Test
NDVI	−0.119	$R^2=0.014$	$P<0.0001$	Important
MSAVI2	0.081	$R^2=0.007$	$P<0.0001$	Important
DSWI	−0.747	$R^2=0.558$	$P<0.0001$	Important
NDGI	−0.769	$R^2=0.592$	$P<0.0001$	Important
MSI	0.743	$R^2=0.552$	$P<0.0001$	Important
BI	0.953	$R^2=0.909$	$P<0.0001$	Important
Crust index	0.662	$R^2=0.438$	$P<0.0001$	Important
TI	−0.606	$R^2=0.367$	$P<0.0001$	Important
Cuirass index	0.942	$R^2=0.887$	$P<0.0001$	Important
Redness index	−0.827	$R^2=0.684$	$P<0.0001$	Important
Colour index	0.662	$R^2=0.438$	$P<0.0001$	Important
GSI	0.685	$R^2=0.470$	$P<0.0001$	Important
NDSI	0.119	$R^2=0.014$	$P<0.0001$	Important

particles. One can understand that the clear soils are much encrusted, dry, formed of particles of coarse size, and characterized by a high brightness. The correlation of the redness and texture indexes with the four vegetation indexes informs on the fact that red soils (hydromorphic and vertisol soils with significant iron content) and sandy soils are mostly present in the vegetated areas and as the soils are battleships vegetation cover decreases. Also, the opposition between NDVI and NDSI reflects the fact that the decline in chlorophyll activity is followed by an increase in soil salinity. Besides, the influence of soil salinity on plant quality and health is also observed by the proximity of NDSI with NDGI and DSWI in the correlation circle.

The first factor informs more about soil-related information and opposes five soil indexes (MSI, BI, crust index, colour index, and GSI) to two vegetation indexes, NDGI and the DSWI. Factor "two" concentrates vegetation information by contrasting the other two main vegetation indexes (NDVI and MSAVI2) with the salinity index (NDSI). Consequently, the main characteristics of the soils derived from the correlations circles between indexes (variables) and the factors are the organic matter content, humidity, and the physical state of the soils for the first factor. The second factor is the state and density of the vegetation cover.

However, the comparison of results obtained between the linear regression and the factorial analysis requires a few remarks. This concerns the consistency of negative correlations between MSAVI2 and NDSI on one hand and positive correlations between MSAVI2 and NDVI on the other. Both are valid for linear regression and descriptive statistics and can then explain why the NDSI is not close to other soil indexes. We notice the low representativeness of the cuirass,

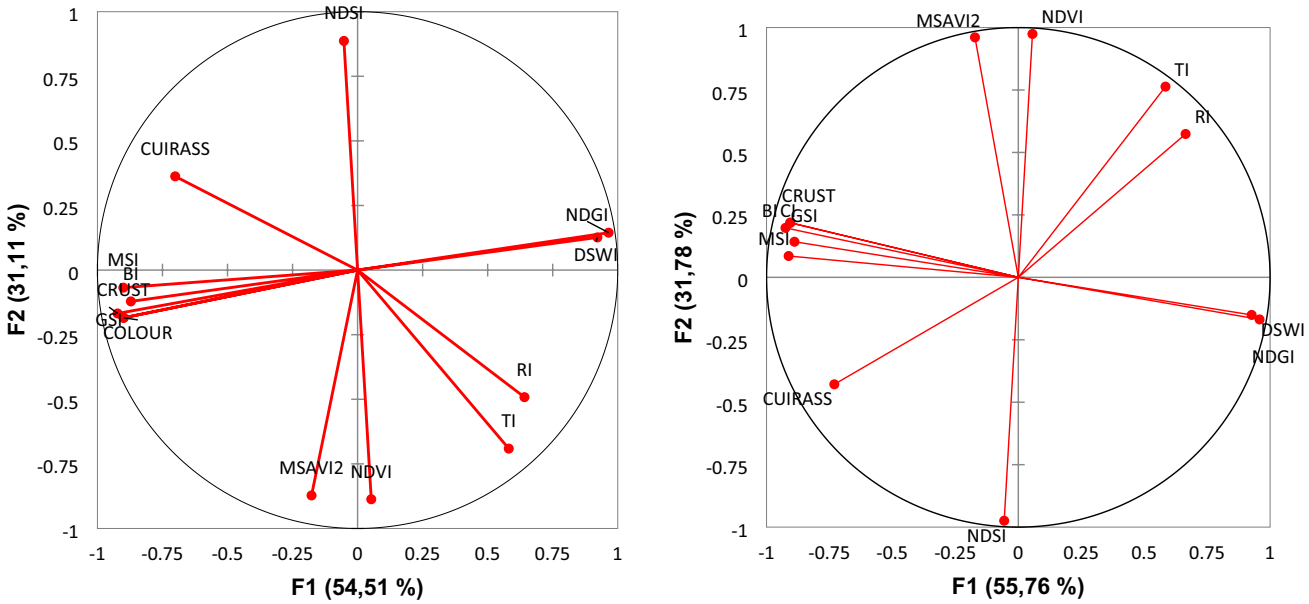

Fig. 6 Correlations between variables and factors

Table 5 Factorial coordinates of indexes

Factorial analysis			Principal component analysis		
	F1	F2		F1	F2
NDVI	0.053	−0.976	NDVI	0.056	0.974
MSAVI2	−0.177	−0.960	MSAVI2	−0.172	0.961
DSWI	0.921	0.139	DSWI	0.927	−0.151
NDGI	0.964	0.159	NDGI	0.958	−0.169
MSI	−0.900	−0.075	MSI	−0.912	0.085
BI	−0.871	−0.133	BI	−0.889	0.143
Crust index	−0.899	−0.203	Crust Index	−0.906	0.219
TI	0.581	−0.761	TI	0.584	0.764
Cuirass index	−0.701	0.399	Cuirass Index	−0.731	−0.428
RI	0.641	−0.541	RI	0.664	0.574
Colour index	−0.899	−0.203	Colour Index	−0.906	0.219
GSI	−0.922	−0.185	GSI	−0.925	0.198
NDSI	−0.053	0.976	NDSI	−0.056	−0.974

redness, and texture indexes and their low correlation with the other indexes. One can also note that in the correlation circle of the PCA, all the variables are far from the centre than they are in the factorial's analysis one. The oppositions between the (variable) indexes remain the same for the PCA as for the factorial analysis only their signs concerning the first factor change.

The Equation Proposed for the Model

The model's equation proposed here is designed to balance all the information obtained from the statistical analysis performed with the indexes and based on previous work approaches. The adopted approach is to weigh the index maps with their coefficient of determination which serves us to highlight the individual contribution of each index to the final map of soil degradation. Also, we consider for each index of its highest values of factorial coordinates obtained through the factorial analysis and the PCA to preserve the best information provided by each of these methods of analysis. This information is combined to compose the following equation:

Results

Map of Exposition Soils Degree to Agents and Degradation Factors

The result of this modelling is a map of exposition soils degree to agents and degradation factors. The potential soil exposition state is classified on the map below in five levels of exposition risk from the "Lower" level to the "Higher" (Table 6). The diversity of land cover explains the nature and the state of the soils, justifies the heterogeneity of the map, and explains the need to have a high number of classes to represent all the levels of exposition risk.

In the absence of field truth data, the different exposition levels are obtained by performing a standard deviation threshold of the image histogram. The standard deviation threshold method allows visualizing how much the attribute values of a class vary compared to the mean, using mean values and standard deviations from the mean.

Table 6 Classification of degradation levels

INDEXES	LOWER ← EXPOSITIONS LEVELS → HIGHER				
NDVI (high to low chlorophyll activity)	0,501 - 0,861	0,285 - 0,501	0,070 - 0,285	-0,145 - 0,070	-0,502 - -0,145
MSAVI2 (high to low vegetation density)	0,425 - 0,925	0,041 - 0,425	-0,342 - 0,041	-0,726 - -0,342	-2,012 - -0,726
DSWI (high to low vegetation water stress)	1,516 - 6,787	1,252 - 1,516	0,988 - 1,252	0,724 - 0,988	0 - 0,724
NDGI (high to low vegetation greenery)	0,039 - 0,370	-0,040 - 0,039	-0,120 - -0,040	-0,200 - -0,120	-0,355 - -0,200
MSI (high to low soil moisture)	0 - 0,657	0,657 - 1,038	1,038 - 1,419	1,419 - 1,622	1,622 - 5,746
BI (low to high soil brightness)	0 - 1394,708	1394,708 - 3280,037	3280,037 - 5165,367	5165,367 - 7050,696	7050,696 - 15541,476
CRUST INDEX (low to high soil crusting)	-0,113 - 0,141	0,141 - 0,205	0,205 - 0,269	0,269 - 0,333	0,333 - 0,607
TI (low to high soil texture)	-0,578 - 0,007	0,007 - 0,078	0,078 - 0,149	0,149 - 0,220	0,220 - 0,428
CUIRASS INDEX (low to high soil cuirass)	-100 - 1703,341	1703,341 - 2439,784	2439,784 - 3176,226	3176,226 - 3912,669	3912,669061 - 366881
RI (high to low soils redness)	2,297e-006 - 1,403e-005	1,731e-006 - 2,297e-006	1,166e-006 - 1,731e-006	6,005e-007 - 1,166e-006	3,366e-009 - 6,005e-007
COLOR INDEX (low to high soil color)	304,098 - 833,117	833,117 - 1588,966	1588,966 - 2344,815	2344,815 - 3100,664	3100,664 - 16903,046
GSI (low to high grain size)	-0,085 - 0,042	0,042 - 0,088	0,088 - 0,133	0,133 - 0,189	0,189 - 0,416
NDSI (low to high soils salinity)	-0,861 - -0,506	-0,506 - -0,291	-0,291 - -0,075	-0,075 - 0,140	0,140 - 0,497

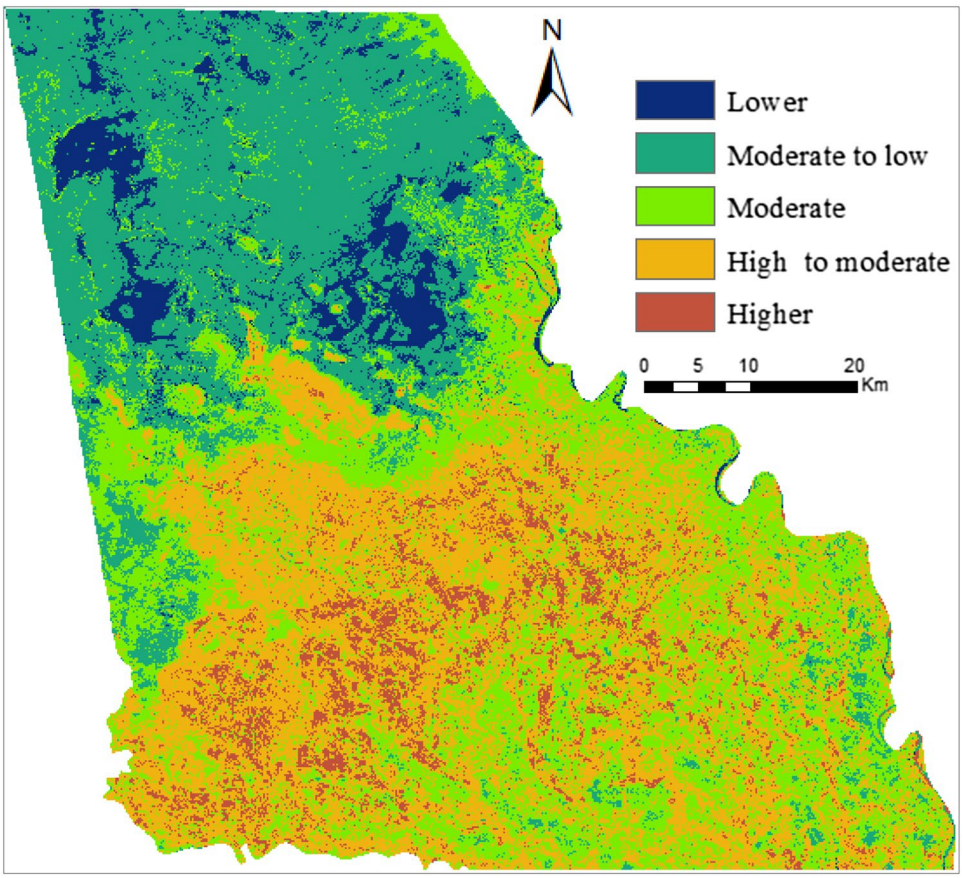

Fig. 7 Map of soils exposition risk to degradation

The “Lower” and “Moderate to low” levels cover the permanent open water areas of the lake, the marshland, and vegetated areas of the immediate shores, a portion of the intermediate shores, and occupy, respectively, 25,214.35 hectares and 130,717.19 hectares (Fig. 7). The "Moderate" level of exposition spreads sparsely over the bare areas

Table 7 Areas of degradation level by index

Level of degradation	NDVI	MSAVI2	DSWI	NDGI	MSI	BI	Crust index	TI	Cuirass index	RI	Colour index	GSI	NDSI
Lower	47,347,15	116,895,5	12,232,53	51,141,44	56,785,6	20,022,04	44,050,68	436,12	34,703,46	12,455,87	42,013,35	9303	45,373,94
Moderate to low	65,710,8	307,273,49	43,821,77	77,454,57	73,233,53	131,856,39	75,252,17	189,099,23	114,273,9	30,834,54	103,409,8	36,286,48	66,397,32
Moderate	310,022,07	9348,34	65,791,41	132,378,86	122,002,11	314,392,22	169,496,11	161,940,13	166,788,83	66,436,07	134,027,8	77,362,84	310,755,2
High to moderate	10,351,48	15,298,62	142,606,5	213,842,42	190,128,43	8575,23	179,840,75	61,304,64	132,939,34	196,491,09	178,548,4	227,054,85	10,428,33
Higher	41,426,95	26,042,49	210,406,3	41,15	32,708,78	15,46	6218,72	62,076,68	26,152,92	168,640,87	16,859,07	124,851,26	41,903,57
TOTAL	474,858,45	474,858,44	474,858,5	474,858,44	474,858,45	474,861,34	474,858,43	474,856,8	474,858,45	474,858,44	474,858,4	474,858,43	474,858,4

of the outer shores and the hinterland over an area of 137,404.34 hectares. The "High to moderate" and "Higher" levels dominate the outer shores and the hinterland. With 152,371.91 hectares, the "High to moderate" level represents the most widespread state of exposition of our study area. The "Higher" level occupies 29,175.73 hectares. The main difficulty now is to be able to identify for each level of exposition the most influential indexes.

The method adopted to answer this concern is inspired by [40], which consists of classifying the indexes by class of degradation and identifying the influence of the indexes according to their spatial distribution by class (Table 7). For the “Higher” level, the top five indexes with the largest spatial distributions are in decreasing order, DSWI (210,406.3 hectares), RI (168,640.87 hectares), GSI (124,851.26 hectares), TI (62,076.68 hectares), and NDSI (41,903.57 hectares). As a result, the “Higher” level is explained by bare soils where vegetation has completely disappeared, the low rate of iron in the soil, the coarse texture of the surface particles, and high salinity. For the “High to moderate” level, the GSI, NDGI, RI, MSI, and crust index with, respectively, 227,054.85 hectares, 213,842.42 hectares, 196,491.09 hectares, 190,128.43 hectares, and 179,840.75 hectares are the most indicative. This means that the soils of this class are also characterized by the coarse texture of the particles on the surface, but also by the weak greenery of the vegetation, an important crusting, and low moisture and iron contents.

The following indexes: brightness (314,392.22 hectares), salinity (310,755.2 hectares), chlorophyll (310,022.07 hectares), crust (169,496.11 hectares), and cuirass (166,788.83 hectares) are the most influential for the "Moderate" level. The soils of this class are clear and salty, weakly covered with vegetation, and compact on their surfaces.

A good vegetation cover of the soil, the fine texture of the soil particles, dark soils, weakly cuirassed, characterizes the "Moderate to low" level that covers the open waters of the lake, marshland areas, and part of the immediate and intermediate shores and which contain organic matter in significant quantities. Indeed, in this class, the MSAVI2 with 307,273.49 hectares is the most widespread index followed by TI 189,099.23 hectares, BI 131,856.39 hectares, cuirass index 114,273.9 hectares, and colour index 103,409.8 hectares. In the "Lower" class, which occupies open water and marshland, the influence of vegetation indexes is the most important (MSAVI2 116,895.5 hectares, NDGI 51,141.44 hectares, NDVI 47,347.15 hectares), the soil moisture is high (MSI 56,785.6 hectares), and their salinity rate is low (the NDSI 45,373.94 hectares).

Table 8 The confusion matrix

Class	Close to nil	Very low	Low	Moderate	High	Total	Commission (%)
Lower	803, 631	147, 710	26, 019	5016	22	982, 398	18, 20
Moderate to low	1, 153, 275	1, 221, 116	356, 131	43, 948	1229	2, 775, 699	56, 01
Moderate	57, 774	481, 539	1, 356, 580	918, 650	70, 123	2, 884, 666	52, 97
High to Moderate	14, 799	97, 778	804, 140	2, 554, 982	770, 687	4, 242, 386	39, 77
Higher	1962	2821	49, 493	492, 528	593, 535	4, 242, 386	47, 95
Total	2, 031, 441	1, 950, 964	2, 592, 363	4, 015, 124	1, 435, 596		
Omission (%)	60, 44	37, 41	47, 67	36, 37	58, 66		

Validation of Results

A confusion matrix was used to validate the results obtained by a comparison with the existing map of the land degradation status of the far north region of Cameroon, provided by [40]. A subset containing the main characteristic of the study area was used, i.e., the permanent open water, the marshland, the immediate shores, the external shores, and the hinterland.

The confusion matrix performed provided the information for verification and accuracy assessment between our results with the ground truth map. The overall accuracy which represents in percent the number of correctly classified values divided by the total numbers of values is 54.3%, and the kappa coefficient which assesses how much better the classification is than a random classification has a value of 40.49% (Table 8).

Conclusion and Discussions

The present work is based on laboratory tests applied to Sentinel 2A satellite images. The purpose was to model the risk of soil degradation in Sahelian regions by combining spectral indexes with statistical analyses. The results are highly correlated to some factors as the phenological season of satellite image acquisition, the quality of the images, the formula of the indexes used, and the applied statistical treatments.

Also, statistical analysis was applied to the resulting image giving on the one hand the correlation and determination coefficients of each index, and on the other hand, the factorial axes which summarize more information. All indexes are considered statistically significant (P-value < 0.0001). The first two factors of PCA and factorial analysis explain, respectively, 87.54% and 85.62% of the common variability of the characteristics measured. The first factor contains the soil information, and the second factor focuses information on vegetation. This final equation of the model is obtained by index weighting with the respective values of the coefficient of determination, which oscillates between 0.007 for the MSAVI2 and 0.909 for the brightness index. Among the most serious levels of degradation, the "High to moderate" level is the most widespread with 15,271.91 hectares, followed by the "Moderate" level with 137,404.34 hectares, and the "Higher" level, which occupies an area of 29,175.73 hectares.

However, we apply our methodology to images of a specific month of the year (April). The challenge now is the adaptation of the model to previous years and other periods of the year. Moreover, the lack of consideration of urban areas is a limit for this work, because the elements that constitute the habitat (example of aluminium roofs) necessarily influence the results of the calculation of certain indexes.

At last, whatever performing decorrelation analysis as a method of unlinking indexes, all of them is calculated on satellite images from the same sensor. Consequently, they have a basic dependent relation because of their origin same spectral characteristics. For this reason, it should be interesting in further analysis to perform the whole analysis on multisource satellite images (SPOT or MODIS), to assess the statistic behaviour and decorrelation, while an index of one source and another of the other source is used as the independent and dependent variable.

Moreover, the method adopted in this study to evaluate the contribution of the different indexes to each degree of degradation brought interesting results. However, the presence on our images of open waters and marshland to a certain extent brings out a new constraint to consider. The low values of vegetation indexes of NDVI and MSAVI2 appear in open water rather than appearing in bare spaces. Without this class of occupation, these two indexes would have better contributed to characterize the classes of strong degradation as the DSWI did. To overcome this difficulty, one of the ways of improving the model will be to classify indexes as functions of the distinct levels of degradation, using the spectral windows obtained from the spectral signature of these indexes.

The imbalance between the number of vegetation index and the number of soil index is to be considered, through a readjustment that will allow integrating new

parameters including climatic like the temperature of the surface, precipitations, albedo, or evapotranspiration. Other elements, such as topography and hydrographic network distribution, are also to be considered.

Acknowledgements The authors are grateful to European Space Agency (ESA) and the Copernicus program for the Sentinel-2 satellite images direct access, the UMR 7300 ESPACE of the French National Research Center (CNRS) for the data processing infrastructure supports. This study was funded by the CNES TOSCA n°12672 BC T54 AIM-CEE program. We thank all those who contributed to this article.

Data availability The data that support the findings of this study are openly available in the Copernicus website at https://scihub.copernicus.eu/, reference number.

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Atman AL, Pradhan B, Saber H, Rahimi A. Desertification sensitivity analysis using MEDALUS model and GIS: a case study of the oases of middle draa valley. Moroc Sens. 2018;18:2230.
2. Ali F, Bouhlassa S. 'Exploitation Des Mesures Magnétiques Dans l'étude de l'état de Stabilité Des Sols: Cas Des Bassins-Versants Abdelali et Markat (Prérif-Maroc)'. Papeles de Geografía 2003;(38):27–40.
3. Allbed A, Kumar L, Aldakheel YY. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: applications in a date palm dominated region. Geoderma. 2014;230–231:1–8.
4. Apan A, Alex H, Stuart P, John M. 'Formulation and Assessment of Narrow-Band Vegetation Indices from EO-1 Hyperion Imagery for Discriminating Sugarcane Disease'. Proc. of the Spatial Sciences Conference, 22–27 September 2003, Canberra; 2003. p. 13.
5. Asfaw E, Suryabhagavan KV, Argaw M. Soil salinity modeling and mapping using remote sensing and GIS: the case of wonji sugar cane irrigation Farm, Ethiopia. J Saudi Soc Agric Sci. 2018;17(3):250–8.
6. Azabdaftari A, Sunar F. Soil salinity mapping using multitemporal landsat data. ISPRS Intern Archiv Photogrammet Remote Sens Spat Inform Sci. 2016;7:3–9.
7. Baccini A. 'Statistique Descriptive Multidimensionnelle'. Publications de l'institut de mathématiques de Toulouse; 2010. p. 33.
8. Bannari A, Huete AR, Morin D, Zagolski F. Effets de la couleur et de la brillance du sol sur les indices de végétation. Int J Remote Sens. 1996;17(10):1885–906.
9. Caloz Régis, Claude Collet 2011 Précis de Télédétection Volume 3 Traitements Numériques d'images de Télédétection. Presses de l'Université du Quebec. 3
10. Chamard PHC, Guenegou MC, Jeannine Lerhun, Levasseur J, Togola M 1991 'Utilisation Des Bandes Spectrales Du Vert et Du Rouge Pour Une Meilleure Évaluation Des Formations Végétales Actives'. In: Congrès AUPELF-UREF, edited by Marie-Françoise Courel. Sherbrooke, Canada. 6
11. Chandana PG, Weerasinghe KDN, Subasinghe S, Pathirana S 2004 'Remote Sensing Approach to Identify Salt-Affected Soils in Hambantota District'. 2nd Academic Sessions, University of Ruhuna. p. 6.
12. Chen Di, Chang N, Xiao J, Zhou Q, Wenbin Wu. Mapping dynamics of soil organic matter in croplands with MODIS data and machine learning algorithms. Sci Total Environ. 2019;669:844–55.
13. Chikhaoui M, Bonn F, Bokoye AI, Merzouk A. A spectral index for land degradation mapping using ASTER data: application to a semi-arid mediterranean catchment. Int J Appl Earth Obs Geoinf. 2005;7(2):140–53.
14. Chikhaoui M, Bonn F, Merzouk A, Lacaze B. Cartographie de la dégradation des sols à l'aide des approches du spectral angle mapper et des indices spectraux en utilisant des données aster. Rev Télédétec. 2007;7(1234):343–57.
15. Eklundh L, Olsson L. Vegetation index trends for the African sahel 1982–1999. Geophys Res Lett. 2003. https://doi.org/10.1029/2002GL016772.
16. Elhag M, Bahrawi JA. Soil salinity mapping and hydrological drought indices assessment in arid environments based on remote sensing techniques. Geosci Instrument Methods Data Syst. 2017;6(1):149–58.
17. Escadafal R, Huete A. 'Etude des propriétés spectrales des sols arides appliquée à l'amélioration des indices de végétation obtenus par télédétection.' Comptes rendus de l'Académie des Sciences. Série 2 Mécanique. 1991;312:1385–91.
18. Escadafal Richard, Bacha S 1996 'Strategy for the dynamic study of desertification'. In: Surveillance des sols dans l'environnement par télédétection et systèmes d'information géographiques edited by Richard Escadafal, MA Mulder, L Thiombiano ORSTOM. Paris, France.
19. Ahmad F. Spectral vegetation indices performance evaluated for cholistan desert. J Geograp Reg Plan. 2012. https://doi.org/10.5897/JGRP11.098.
20. Gao Y, Gao J, Wang J, Wang S, Li Q, Zhai S, Zhou Ya. Estimating the biomass of unevenly distributed aquatic vegetation in a lake using the normalized water-adjusted vegetation index and scale transformation method. Sci Total Environ. 2017;601–602:998–1007.
21. Gbetkom PG, Gadal S, El Aboudi A, Mfondoum AHN, Badamassi MBM. Mapping change detection of LULC on the cameroonian shores of lake chad and its hinterland through an inter-seasonal and multisensor approach. Intern J Advan Remote Sens GIS. 2018;7(1):2835–49.
22. Gérard B, Richard E, Delphine F, Anne-Thérèse H-NN. 'La Télédetection : Un Outil Pour Le Suivi et l'évaluation de La Désertification'. Montpelli (FRA); Montpellier: CSFD; Agropolis, p. 36 (Les Dossiers Thématiques - CFSD ; 2); 2005.
23. GIZ 2015 'Audit Environnemental Conjoint Sur l'Assèchement Du Lac Tchad'. Afrique du Sud.
24. Gorji T, Tanik A, Sertel E. Soil salinity prediction, monitoring and mapping using modern technologies. Procedia Earth Planetary Sci. 2015;15:507–12.
25. Guerrien M. L'intérêt de l'analyse en composantes principales (ACP) pour la recherche en sciences sociales: présentation à partir d'une étude sur le mexique. Cahiers Des Amériques Latines,. 2003;43(July):181–92.
26. Houssa R, Pion J-C, Yésou H. Effects of granulometric and mineralogical composition on spectral reflectance of soils in a sahelian area. ISPRS J Photogramm Remote Sens. 1996;51(6):284–98.
27. El Jazouli A, Barakat A, Khellouk R, Rais J, El Baghdadi M. Remote sensing and GIS techniques for prediction of land use land cover change effects on soil erosion in the high basin of the

oum er rbia river (Morocco). Remote Sens Appl Soc Environ. 2019;13:361–74.
28. Karnieli A. Development and implementation of spectral crust index over dune sands. Int J Remote Sens. 1997;18(6):1207–20.
29. Kauth RJ, Thomas GS. The tasselled cap—A graphic description of the spectral-temporal development of agricultural crops as seen by LANDSAT. LARS Symposia. LARS Symposia. Paper 159. 1976. p. 13. http://docs.lib.purdue.edu/lars_symp/159
30. Khan NM, Rastoskuev VV, Sato Y, Shiozawa S. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. Agric Water Manag. 2005;77(1–3):96–109.
31. Li H, Zhao C, Yang G, Feng H. Variations in crop variables within wheat canopies and responses of canopy spectral characteristics and derived vegetation indices to different vertical leaf layers and spikes. Remote Sens Environ. 2015;169:358–74.
32. Li X, Zhang Y, Bao Y, Luo J, Jin X, Xingang Xu, Song X, Yang G. Exploring the best hyperspectral features for LAI estimation using partial least squares regression. Remote Sens. 2014;6(7):6221–41.
33. Netto Madeira J. Etude quantitative des relations constituants minéralogiques-réflectance diffuse des latosols brésiliens: application à l'utilisation pédologique des données satellitaires TM (région de Brasilia). Bondy: ORSTOM. Centre IRD de Bondy; 1991.
34. Maimouni S, Bannari A, El-Harti A, El-Ghmari A. Potentiels et limites des indices spectraux pour caractériser la dégradation des sols en milieu semi-aride. Can J Remote Sens. 2011;37(3):285–301.
35. Maimouni S, Abderrazak B. 'Cartographie de La Dégradation Des Sols En Milieu Semi-Aride'. 32e Symposium canadien sur la télédétection et du 14e Congrès de l'AQT. 2011. p. 7.
36. Mandal Umesh K. Spectral color indices based geospatial modeling of soil organic matter in Chitwan District, Nepal. Int Arch Photogram Rem Sens Spat Inf Sci. 2016;XLI-B2:43–8.
37. Martín-Sotoca JJ, Saa-Requejo A, Borondo J, Tarquis AM. Singularity maps applied to a vegetation index. Biosys Eng. 2018;168(April):42–53.
38. Mathieu R, Pouget M, Cervelle B, Escadafal R. Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment. Remote Sens Environ. 1998;66(1):17–28.
39. Narmada K, Gobinath K, Bhaskaran G. Monitoring and evaluation of soil salinity in terms of spectral response using geoinformatics in Cuddalore environs. Intern J Geomat Geosci. 2015;5(4):536–43.
40. Ngandam MA, Homère JE, Nongsi BK, Moto FAM, Deussieu FGN. Assessment of land degradation status and its impact in arid and semi-arid areas by correlating spectral and principal component analysis neo-bands. Intern J Advan Remote Sens GIS. 2016;5(1):1539–60.
41. Okaingni J-C, Kouamé KF, Martin A. Cartographie des cuirasses dans les formations volcano- sédimentaires de la zone d'anikro-kadiokro (côte d'ivoire) à l'aide de la théorie des fonctions de croyance. Revue Télédétect. 2010;9(1):19–32.
42. de Oliveira Pedro D, Sato MK, Rodrigues S, de Lima HV. S-index and soybean root growth in different soil textural classes. Rev Brasileira de Engenharia Agrícola e Ambiental. 2016;20(4):329–66.
43. Palm R. Les méthodes d'analyse factorielle : principes et applications. Notes de Statistique et d'Informatique. 1993;1:36.
44. Pandey PC, Rani M, Srivastava PK, Sharma LK, Nathawat MS. Land degradation severity assessment with sand encroachment in an ecologically fragile arid environment: a geospatial perspective. QScience Connect,. 2013;2013:43.
45. Pang G, Wang X, Yang M. Using the NDVI to identify variations in, and responses of, vegetation to climate change on the tibetan plateau from 1982 to 2012. Quatern Int. 2017;444:87–96.
46. 'Plan d'Action National de Lutte Contre La Désertification (PAN/LCD)'. 2006. République du Cameroun, Ministère de l'environnement et de la protection de la nature. 97
47. Pouchin T. Cours de Télédétection. Université Le Havre France. In: Guerinial A (ed.) Analyse spatio-temporelle par télédétection de la région de Djelfa. Ecole Nationale Supérieure Agronomique El Harrach-Alger. Mémoire, 2012. 2001. p. 44. http://dspace.ensa.dz:8080/jspui/bitstream/123456789/245/1/gueriniai_a.pdf.
48. Pu R. An exploratory analysis of in situ hyperspectral data for broadleaf species recognition. In: International Conference on Photogrammetry, Remote Sensing, and Spatial Information Sciences. Beijing, China. 2008. p. 6.
49. Qi J, Chehbouni A, Huete AR, Kerr YH, Sorooshian S. A modified soil adjusted vegetation index. Remote Sens Environ. 1994;48(2):119–26.
50. Rallo G, Minacapilli M, Ciraolo G, Provenzano G. Detecting crop water status in mature olive groves using vegetation spectral measurements. Biosys Eng. 2014;128:52–68.
51. Ray SS, Singh JP, Das G, Panigrahy S. Use of high resolution remote sensing data for generating site-specific soil management plan. Int Arch Photogramm Remote Sens Spat Inf Syst B. 2014;35(7):6.
52. Renard KG, Foster GR, Weesies GA, McCool DK, Yoder DC. Predicting soil erosion by water: a guide to conservation planning with the revised universal soil loss equation (RUSLE). DC, USA: US Government Printing Office; 1997. p. 703.
53. Romero M, Luo Y, Baofeng Su, Fuentes S. Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. Comput Electron Agric. 2018;147:109–17.
54. Rouse JW, Haas RH, Schell JA, Deering DW. Monitoring vegetation systems in the great plains with ERTS. NASA Spec Publ. 1973;351(1):309–17.
55. Seignobos Christian, Olivier Iyébi-Mandjek 2000 'Atlas de la province Extrême-Nord Cameroun'. Paris : Yaoundé: Institut de recherche pour le développement ; République de Cameroun, Ministère de la recherche scientifique et technique, Institut national de cartographie.
56. StéphaneKoff A, Fora AA, Elbelrhiti H. Cartographie de l'état du couvert végétal du nord de la côte d'ivoire à partir d'images satellites: exemple de la zone de korhogo. Europ Sci J ESJ. 2016;12(29):204.
57. Sun H, Li M, Li D. The vegetation classification in coal mine overburden dump using canopy spectral reflectance. Comput Electron Agric. 2011;75(1):176–80.
58. Symeonakis E, Drake N. Monitoring desertification and land degradation over sub-saharan Africa. Int J Remote Sens. 2004;25(3):573–92.
59. Symeonakis E, Drake N. 10 daily soil erosion modelling over sub-saharan Africa. Environ Monit Assess. 2010;161(1–4):369–87.
60. Welikhe P, Quansah JE, Fall S, McElhenney W. Estimation of soil moisture percentage using LANDSAT-based moisture stress index. J Remote Sens GIS. 2017. https://doi.org/10.4172/2469-4134.1000200.
61. Wischmeier WH, Smith DD. Predicting rainfall erosion losses: a guide to conservation planning. Science and Education Administration: Department of Agriculture; 1978.
62. Xiao J, Shen Y, Tateishi R, Bayaer W. Development of topsoil grain size index for monitoring desertification in arid land using remote sensing. Int J Remote Sens. 2006;27(12):2411–22.
63. Xiao J, Yanjun S, Tateishi R. Mapping soil degradation by topsoil grain size using MODIS data. Chiba University. Center for Environmental Remote Sensing. 2014. p. 8.
64. Yongnian Zeng, Zhaodong Feng, Nanping Xiang 2004 'Assessment of Soil Moisture Using Landsat ETM+ Temperature/Vegetation Index in Semiarid Environment'. In: IEEE

International IEEE International IEEE. International Geoscience and Remote Sensing Symposium. IEEE. Anchorage, USA. 4306