

Article

Spectral Diversity Metrics for Detecting Oil Pollution Effects on Biodiversity in the Niger Delta

Nkeiruka Nneti Onyia ^{1,*} , Heiko Balzter ^{1,2}  and Juan Carlos Berrio ¹ 

¹ School of Geography, Geology and Environment, University of Leicester, University Road, Leicester LE1 7RH, UK; jcb34@leicester.ac.uk

² National Centre for Earth Observation, University of Leicester, University Road, Leicester LE1 7RH, UK; hb91@leicester.ac.uk

* Correspondence: nno4@le.ac.uk; Tel.: +44-79-563-610-73

Received: 27 July 2019; Accepted: 11 November 2019; Published: 14 November 2019



Abstract: Biodiversity monitoring in the Niger delta has become pertinent in view of the incessant spillages from oil production activities and the socio-economic impact of these spillages on the inhabitants who depend on the resources for their livelihood. Conventional methods of post-impact assessments are expensive, time consuming, and cause damage to the environment, as they often require the removal of affected samples/specimens for laboratory analysis. Remote sensing offers the opportunity to track biodiversity changes from space while using the spectral variability hypothesis (SVH). The SVH proposes that the species diversity of a sampled area is linearly correlated with the variability of spectral reflectance of the area. Several authors have tested the SVH on various land cover types and spatial scales; however, the present study evaluated the validity of the SVH against the backdrop of oil pollution impact on biodiversity while using vascular plant species as surrogates. Species richness and diversity indices were computed from vegetation data collected from polluted and non-polluted transects. Spectral metrics that were derived from Sentinel 2 bands and broadband vegetation indices (BVIs) using various algorithms, including averages, spread, dimension reduction, and so on, were assessed for their ability to estimate vascular plants species richness and diversity. The results showed significant differences in vegetation characteristics of polluted and control transects ($H = 76.05$, p -value = <0.05 for abundance and $H = 170.03$, p -value < 0.05 for richness). Spectral diversity metrics correlated negatively with species data on polluted transects and positively on control transects. The metrics computed using Sentinel 2A bands and vegetation indices proved to be sensitive to changes in vegetation characteristics following oil pollution. The most robust relationship was observed between the metrics and indices on control transects, whereas the weakest relationships were observed on polluted transects. Index-wise, the Simpson's diversity index regressed better with spectral metrics ($R^2 > 0.5$), whereas the Chao-1 richness index regressed the least ($R^2 < 0.5$). The strength of the relationship resulted in successfully estimating species richness and diversity values of investigated transects, thereby enhancing biodiversity monitoring over time and space.

Keywords: Spectral Variability Hypothesis (SVH); Sentinel 2; biodiversity; spectral metrics; vegetation indices; diversity indices

1. Introduction

The Spectral Variability Hypothesis (SVH) that was proposed by Palmer et al. [1] asserts that the spatial heterogeneity of plant species positively correlates with spectral diversity of remotely sensed images. Spectral diversity, on the one hand, is determined by the extent of variation in the reflectance values of corrected images. Plants, like other materials on the earth surface, uniquely

reflect radiation from the sun for different species. The shape of the spectrum is dependent on the optical properties of the plants, which in turn are controlled by biophysical and biochemical factors, such as cellular structure, leaf anatomy, pigment concentrations, and moisture content. According to Clevers et al. [2], the control of leaf reflectance is by pigments (chlorophyll a and b, β -carotene, and so on) in the visible (VIS); the cellular structure in the near-infrared (NIR) and moisture content in the short-wave infrared (SWIR) regions of the electromagnetic spectrum. Variations in these properties that are induced by plant response to environmental stress, such as oil pollution, manifest in the spectral signatures acquired by satellite sensors in orbit. Plants response to environmental stress has been well-studied [3], as well as the adverse effect of such stressors to plant productivity [4]. Environmental stressors that affect plants also result in biodiversity loss globally [5,6].

Hydrocarbons in the soil displace oxygen and nutrients and adversely affect plant growth and productivity [7–9]. Zhang et al. [10] investigated the effect of lubricating oil pollution on mangrove species in Hong Kong and reported that the hydrocarbons induced the formation of excessive free radicals in plant cells. These radicals caused oxidative stress by impairing biological processes that enhance growth and productivity in plant seedlings. In their work at a leakage site in Central Bohemia, Czechoslovakia (present-day Czech Republic) Pyšek and Pyšek [6] noted a decrease in the growth and number of individual plants that are present on contaminated sites.

Species diversity is an ecosystem status indicator according to the Group on Earth Observation (GEO BON), the World Climate Research Programme (WRCP), the International Geosphere-Biosphere Program (IGBP), and the Committee on Earth Observation systems (CEOS) [11]. Past studies show links between species diversity and ecosystem productivity, such that loss of the former threatens the goods and services derived from the latter [12–16]. Similarly, Pereira and Cooper [17] observed that species diversity determines the ecosystem functioning and type of ecosystem services provided at the local, regional, or global scale. Effort to minimize or halt the rapid and significant loss of biodiversity worldwide led to the development of initiatives and frameworks (such as the Convention on Biological Diversity, CBD) and gauging the effectiveness of these policies require timely, standardized, and scalable information on the status and trends of biodiversity [18,19]. Remote sensing offers great potential for the achievement of these goals and provides the basis for the development of essential biodiversity variables by GEO BON [19]. Remote sensing is used in biodiversity monitoring to estimate plant species diversity while using very high spectral and spatial resolution images that are analysed for spectral diversity [19–22]. Hence, the SVH represents a valuable resource that integrates spectral and field data for detecting various effects of oil pollution on the ecosystem.

The spectral variability hypothesis (SVH) was tested in this study to determine the suitability of spectral diversity metrics that are derived from Sentinel 2A data for biodiversity monitoring in oil-polluted locations. To achieve this, the following questions were tackled

1. Is there any difference in the reflectance spectra of polluted and non-polluted vegetation?
2. Is there a significant relationship between spectral diversity metrics and field measured variables?
3. Is the application of the SVH valid for oil polluted locations?

2. Materials and Methods

2.1. Study Area

Rivers State is located in Nigeria between latitude 4.75° North and longitude 6.83° East in the eastern part of the Niger Delta, on the ocean ward extension of the Benue trough (Figure 1). It occupies an area approximately 11,077 km² of the delta described as constructive and mainly fed with sediments from the heavily-laden (about 330,000 cm³/annum) Niger River originating from Guinea travelling through Mali, Niger, Benin Republic, and Northern Nigeria [23]. The Delta also receives sediments from the Benue River, albeit to a lesser degree. Rivers State is bounded on the South by the Atlantic Ocean, to the North by Imo, Abia, and Anambra States, to the East by Akwa Ibom State, and to the West by Bayelsa and Delta States. It receives some of the highest rainfall in the world of up to 3000 mm

annually [24]. The temperature ranges from 20° to 30° C during the day. The Intertropical Front wind originates from the meeting point of humid air masses of the Gulf of Guinea and dry air masses from the north continuously blow over the delta, which results in high levels of humidity of up to 75% [25]. The area is also characterised by high cloud cover, which invariably affects the quality of satellite data obtained at certain times of the year [24].

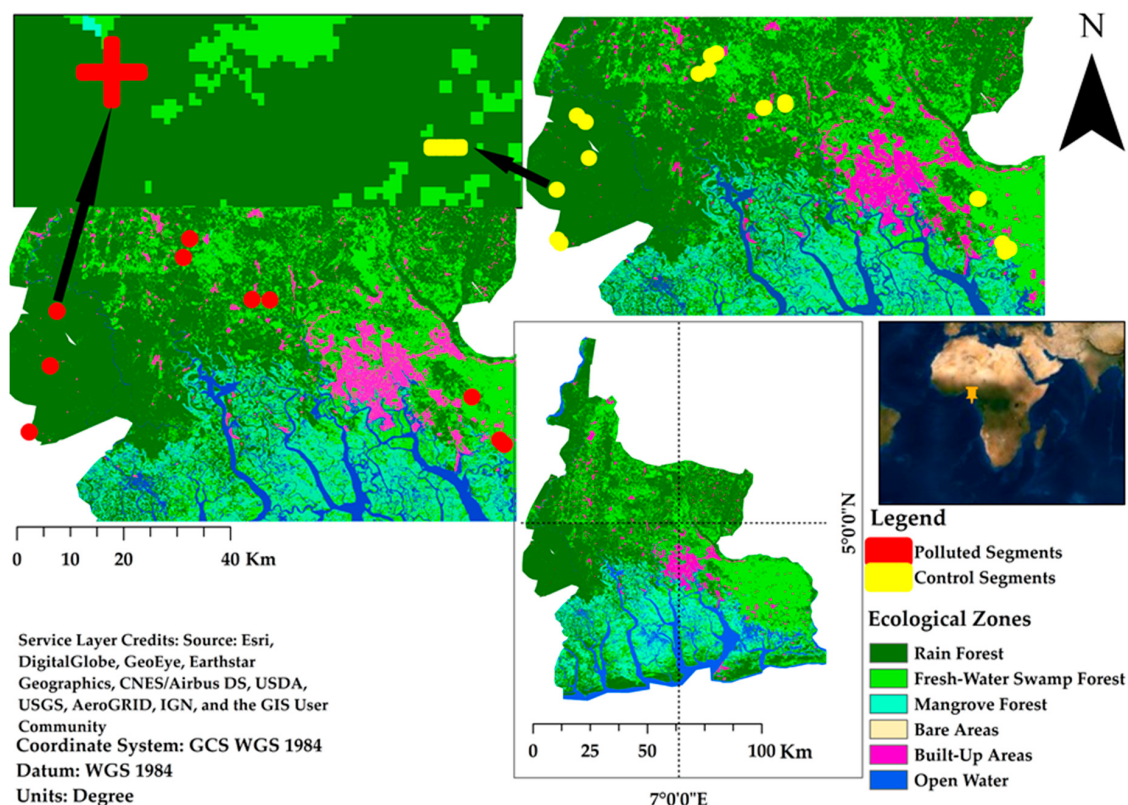


Figure 1. Map of Rivers State showing the location of investigated spill transects (Polluted and non-polluted i.e., control), ecological zones, and waterbodies (rivers and creeks).

In line with the rest of the Delta geomorphology, Rivers State consists of alluvial deposits of sands, silts, and clays deposited during the late Miocene-Pliocene times. Coarse sands and gravels underlie parts of the area while fine sands and clays underlie other areas [26]. The landscape of the area is generally flat with altitude that ranges between 6 to 14m above sea level. The investigated transects located in various local government areas fall in the coastal plain and freshwater ecological zone populated by forest tree species, mangrove, palms, shrubs, ferns, lianas, and so on; however, the rainforest is degraded as observed during the field work. Vegetation around spill transects is variable with evidence of fire occurrence within 30 m radius of the spill epicentre at one of the polluted locations in Kporghor. Other features include untarred roads, which are part of the oil companies' right of way (ROW), vegetated land (natural and farmed), and bare soil. Records from the Nigerian Oil Spill Monitor website [27] (<https://oilspillmonitor.ng>) indicate that the spills occurred between July to December 2015 with sabotage being the leading cause. The sabotage of oil pipelines is the typical method by which crude oil is illegally extracted or diverted. It involves exploding dynamites near pipelines, loosening valves, cutting the pipes, drilling holes in the pipes to fit taps, and applying corrosive substances all in the bid to access the crude oil [28]. The extracted crude is either sold in the black market or locally refined for personal use. Estimated spill volumes reported the range from 46 barrels to over 5000 barrels. Hence, between July and December, 2015 (about six months) over 10,641 barrels of crude oil were spilt into the vulnerable ecosystem of Rivers State in the Niger Delta region of Nigeria.

2.2. Soil Sampling and Vegetation Survey

Vegetation sampling involved the line-intercept method that was discussed in Cummings and Smith [29]. The line intercept method was sufficient for capturing field data to determine species diversity and abundance due to the unique circumstances prevailing at the study area during this campaign [30]. Ten oil spill locations in the Rivers State of Nigeria were investigated. Most of the spills emanated from damaged pipelines that crisscross the landscape of the Niger Delta; however, a few occurred at oil well wellhead locations that were inaccessible; hence, survey was limited to affected and accessible pipelines. Four transects 100 m long transverse the polluted locations in the four cardinal directions starting from the spill epicentre were labelled A, B, C, and D, respectively. The spill epicentre was marked by a 20 × 20 m quadrant from where transects were measured and labelled SS0. Two unpolluted (control) transects that were located within the same towns and ecological zones to ensure ecological similarity, but at least 2 km from the polluted transects, were labelled control 1 and control 2. Control transects were randomly located in areas that have no record of oil spill based on the data available from the Nigerian oil spill monitor website. Polluted transects were subdivided into five segments of 20 m length labelled SS1 to SS5, corresponding to increasing distance from the spill epicentre (SS0). Similarly, control transects were also subdivided into five segments. For each segment, a composite soil sample, species richness and diversity (alpha) values, and vegetation abundance based on the number of individuals observed in the segment were recorded. Abundance is a unit-less measure and it was calculated for each segment as the ratio of the number to occurrence frequency of individual plants [31]. The soil samples were analysed in the laboratory for total petroleum hydrocarbon (TPH) concentration. However, to correct for spatial autocorrelation effects, analysis of data involved alternate segments on each transect including the SS0s.

Inventory of all vascular plant species present on transects was done in February and March 2016 and in May and June 2017 with the aid of prepared tally sheets and local experts. Tally sheets listing indigenous species and photographs were taken to the field to help in species identification with consideration given to the type and shape of leaf, margin, apex, and base of each species; as well as the arrangement of leaves and leaflets on the petioles. Previous reports of common species in the Niger Delta area, such as Ubom [32] and Agbagwa and Ekeke [33] provided material for tally sheets. A deliberate effort was made to minimise any damage to vegetation during the field campaign, hence photographs of unknown plants were taken to the Herbariums of the Forestry Research Institute of Nigeria, Umuahia, and the Michael Okpara University of Agriculture, Umudike both in Abia State, Nigeria for identification.

2.3. Sentinel-2A Image Acquisition and Processing (S2AD)

The Multi-Spectral Imager (MSI) sensor on board the Sentinel 2A satellite acquired the images that were used in this analysis. The Sentinel-2A satellite launched on the 23rd of June 2015 is one of the fleets of satellites owned by the European Commission (EC), in partnership with the European Space Agency (ESA). It is designed to provide imagery that supports environmental monitoring under the EC/ESA's Copernicus programme. The data from this satellite are open and freely downloadable from the Copernicus website. The MSI sensor acquires images in 13 spectral bands with wavelengths ranging from 443–2190 nm. The spectral bands include three visible (Red, Green, and Blue), one near infrared (NIR), and three short-wave infrared (SWIR) bands. As an added advantage, the MSI also has three red-edge bands, which are useful for differentiating crop types and detecting vegetation stress and one coastal aerosol band useful for atmospheric correction (<http://www.esa.int/Copernicus/Sentinel-2>). However, only eight bands relevant for vegetation analysis were used for the present study. These were the visible (bands 2, 3, and 4), NIR (band 8) and red-edge bands (bands 5, 6, 7, and 8A). Table 1 shows the bands, bandwidths, central wavelengths, and spatial resolution of the Sentinel 2A image relevant to this study.

Table 1. Spectral and spatial resolution of Sentinel-2A image.

Sentinel-2A Bands	Band Number	Bandwidth (nm)	Central Wavelength (nm)	Resolution (m)
Blue	2	65	490	10
Green	3	35	560	10
Red	4	30	665	10
Vegetation Red Edge	5	15	705	20
Vegetation Red Edge	6	15	740	20
Vegetation Red Edge	7	20	783	20
NIR	8	115	842	10
Vegetation Red Edge	8A	20	865	20

Spatially, Sentinel-2A images have a swath width of 290km and a resolution of 10 m (VNIR), 20 m (Red-edge and SWIR), and 60 m (atmospheric correction bands). These are some of the best spatial resolutions for freely available satellite data. Six level 1C processed images (geometric and radiometrically corrected) were downloaded from the Copernicus Services Data Hub (<https://cophub.copernicus.eu/>) to cover the study area. These images were acquired between 29 December 2016 and 5 January 2017 and downloaded as 100 by 100 km² granules were selected based on very low cloudy pixel percentage. Each image was atmospherically corrected while using the Sen2cor plugin tool to remove the atmospheric effects from sentinel 2A level 1C images. The processor computes surface reflectance from the top of the atmosphere reflectance values in the level 1C images. Details of the procedure are documented and are available at the ESA SNAP website (<http://www.step.esa.int/main/third-party-plugins-2/sen2cor>). Further analysis performed on the images included band analysis, vegetation index computation, and derivation of spectral metrics. These were completed before basic geo-processing procedures, such as resampling, mosaicking, and clipping were done, in order to retain as much original information in the pixel texture as possible. Rocchini et al. [11] observed that the image analysis that involves smoothing processes could cause a loss of vital information. Resampling was undertaken to downscale the pixel resolution of the various bands to 10 m to achieve uniformity in the results. This procedure was necessary for also performing further image processing on the ESA SNAP platform. Upsampling (resampling to a higher resolution) that was performed in SNAP used the bilinear interpolation method, which is highly recommended for satellite data.

2.4. Scale Matching Satellite Image and Study Area

Sentinel 2A image was spatially scaled to match the sampling units of the field survey. Rocchini et al. [34] suggested that appropriately scaling imagery resolution with species data was essential for implementing the SVH. Similarly, Small [35] and Rocchini [36] noted that matching the field sampling units with the spatial resolution of an image will enhance the detection of sub-pixel variability and strengthen the relationship between species diversity and spectral variability. Additionally, Turner et al. [37] and Chen and Henebry [38] suggested that the calculation of spectral variability is enhanced when several pixels cover the spatial dimension of sampling units.

Vector polygons of each segment on the investigated transects were created in ENVI 5.3. Each segment of 20 m corresponded to 2 × 10 m pixels of the Sentinel 2A image; however, a 2 × 2-pixel window was used in this analysis to incorporate information from the surrounding area. Although the total area of the pixel window was more than the sampling units (segments), this was not considered to be limitation, since incorporating spectral information of landscape surrounding sampling units improve the performance of models linking species to spectral diversity [34,39]. In total, 210 × 4 pixels

(equivalent to an area of 84,000 square meters) were used to test the spectral variability hypothesis in this study.

2.5. Spectral Diversity Metrics

Spectral diversity metrics are spatial quantifications of the diversity of satellite imagery (Warren et al. 2014). Metrics were computed from Sentinel 2A bands and from known vegetation indices that were computed from the Sentinel 2A data. Vegetation reflectance in the 400 to 2500 nm region of the electromagnetic spectrum provides valuable information for the remote sensing of vegetation. Previous research using both in situ and satellite spectral data show that various plant physiological parameters control both leaf and canopy reflectance at different regions of the spectrum. For instance, Gitelson [40] and Wu [41] showed that chlorophyll pigments strongly influence reflectance in the red and blue spectral regions. Diversity metrics were computed from individual bands of the Sentinel 2 image in order to maximize the information inherent in vegetation spectrum. This is because according to Mielke, Schaffer and Schilling [42], differences in leaf internal structure and presence of other pigments affect leaf reflectance of different species at similar wavelengths. Hence, band-based diversity metrics are expected to reveal differences in the species composition of the study area in line with the SVH.

Furthermore, vegetation indices are derived from the application of mathematical expressions on the spectral reflectance of plant materials. They are functions of the reflectance in visible and near-infrared (NIR) spectral bands and they are designed to enhance sensitivity to the parameter of interest, while minimising atmospheric interferences. Most indices are derived from a combination of visible and NIR reflectance depending on the objective of the study [43]. Based on the nature of vegetation indices, it is expected that the spectral diversity of selected indices might be a reflection of the diversity of the measured vegetation. Both sets of indices were compared on their predictive performance in models estimating species diversity in the study area.

2.6. Spectral Diversity Metrics from Bands

The study only utilized eight bands out of the original 13 bands of the Sentinel-2A imagery, these were bands 2 = blue; 3 = green; 4 = red; 5–7, 8A = red edge and 8 = near infrared (NIR). A vector layer of polygons representing segments of investigated transects was overlaid on the pre-processed satellite imagery. For each segment measuring 20m in length, four pixels were identified to encompass it and each pixel was treated as a distinct species to determine the spectral variability of the locations. The reflectance of individual Sentinel-2A bands was then extracted from these four pixels and used in computing the spectral diversity of segments based on various approaches that were adopted from literature. Most procedures involved the center and spread of the band reflectance values from the corresponding pixels, others included the use of information theory (Shannon index) and probability (Simpson index). Metrics computed from the mean, standard deviation, Shannon and Simpson's indices of the pixels followed the method outlined in Warren et al. [44]. Two additional metrics that were defined as spectral heterogeneity (SH) and quartile-based coefficient of variation (QCV) derived from the methods of Hall et al. [21] and Heumann, Hackett and Monfils [20], respectively, were computed from the Sentinel 2A bands. SH is the mean difference between the mean of each 2×2 -pixel window overlaying a segment and the mean of all 2×2 -pixel windows on each transect (overlaying five segments in total). QCV is a non-parametric approach that measures the dispersion of data around a center (median value) by taking the ratio of the interquartile range to the median of the data set (IQR/Median). Further spectral metrics computed from the original 8 Sentinel-2A bands include those that were obtained by principal component analysis (PCA), which is a mathematical algorithm that transforms high dimensional and correlated data to lower dimension and uncorrelated components. Data reduction involves identifying the direction (eigenvectors) of the most variance in the data [45]. These statistics were calculated while using reflectance values of eight bands from the four pixels overlaying the investigated segments. In total, 56 band indices were derived from the eight Sentinel-2A bands utilized. Table 2 presents a list of derived metrics.

Table 2. Spectral indices derived from the Sentinel-2A bands used in the study. Band reflectance of four pixels overlaying each segment provided data for this analysis.

Band	Statistic	Index
2 (Blue)	Mean	M2
	Standard deviation	Sd2
	Spectral heterogeneity	Sh2
	Quartile coefficient of variation	QCV2
	Shannon's Index	H2
	Simpson's Index	D2
	Sorenson's Similarity Index	SI2
3 (Green)	Mean	M3
	Standard deviation	Sd3
	Spectral heterogeneity	Sh3
	Quartile coefficient of variation	QCV3
	Shannon's Index	H3
	Simpson's Index	D3
	Sorenson's Similarity Index	SI3
4 (Red)	Mean	M4
	Standard deviation	Sd4
	Spectral heterogeneity	Sh4
	Quartile coefficient of variation	QCV4
	Shannon's Index	H4
	Simpson's Index	D4
	Sorenson's Similarity Index	SI4
5–7, 8A (Red Edge)	Mean	M5–M7, M8A
	Standard deviation	Sd5–Sd7, Sd8A
	Spectral heterogeneity	Sh5–Sh7, Sh8A
	Quartile coefficient of variation	QCV5–QCV7, QCV8A
	Shannon's Index	H5–H7, H8A
	Simpson's Index	D5–D7, D8A
	Sorenson's Similarity Index	SI5–SI7, SI8A
8 (NIR)	Mean	M8
	Standard deviation	Sd8
	Spectral heterogeneity	Sh8
	Quartile coefficient of variation	QCV8
	Shannon's Index	H8
	Simpson's Index	D8
	Sorenson's Similarity Index	SI8
PC1	Scores	PC1
PC2	Scores	PC2
PC3	Scores	PC3

2.7. Spectral Diversity Metrics from Vegetation Indices (VIs)

Table 3 lists vegetation indices that are useful for detecting the chlorophyll content, primary productivity, and vegetation stress in plants extracted from the original bands of the Sentinel-2A image. Metrics that were derived from VIs include the mean, standard deviation, Shannon's and Simpson's diversity indices, which were used in further analyses.

Table 3. Summary of selected vegetation indices used in evaluating the spectral variation hypothesis (SVH).

Parameter	Index	Formula	Reference
Chlorophyll Content	Canopy Chlorophyll Index (CCI)	$\frac{Band8 - Band5}{Band8 + Band4}$	[46]
	Normalised Difference Vegetation Index (NDVI)	$\frac{Band8 - Band4}{Band8 + Band4}$	[47,48]
	Red Edge Position 2	$\frac{Band5 - Band4}{Band5 + Band4}$	[49]
Primary Productivity	Soil Adjusted Vegetation Index (SAVI)	$\frac{Band8 - Band4}{Band8 + Band4 + L} (1 + L)$	[50]
Leaf Pigments	Anthocyanin Reflectance Index (ARI)	$\frac{1}{Band3} - \frac{1}{Band5}$	[51]
	Structure Insensitive Pigment Index (SIPI)	$\frac{Band8 - Band1}{Band8 - Band4}$	[52]

2.8. Statistical Analysis

The analysis of vegetation data utilized the Paleontological Statistics Software Package, (PAST) to compute the species richness and diversity of the study area and ecological packages *labdsv* and *vegan* to compute the phytosociological characteristics of vegetation in R-studio. Spectral diversity metrics of segments across the entire study area were explored to identify the outliers and the probability distribution while using the Anderson–Darling normality test. The results showed that the datasets did not fulfil the assumptions of a normal distribution; hence, the application of non-parametric analytical methods. The Kruskal–Wallis one-way analysis of variance by ranks was employed to test the null hypothesis of no difference in median values of more than two independent samples [53]. It compared soil TPH and vegetation abundance data from segments of polluted and control transects for all locations investigated, and significant results from the omnibus test were subjected to pairwise multiple-comparisons of mean rank sums while using the Dunn's test. Dunn's test identified which samples differed significantly. The Bonferroni correction procedure adjusted the *p*-values, which controlled the family-wise error that might lead to false discoveries. The Bonferroni adjustment divides the overall alpha (0.05) by the total number of multiple tests [54].

The correlation of each spectral metric with each species diversity measure identified the most sensitive metrics to oil pollution. Following the Spearman's Rank Correlation (SRC) of the data, metrics with large coefficient values ($r > \pm 0.2$) were selected and further tested for significance at $\alpha = 0.05$ while using the Student *t*-test. A *p*-value of less than 0.05 implies that there is sufficient evidence to conclude that there is a significant linear relationship between the spectral metric and the particular field-measured variable and that the relationship is replicable. Selected spectral metrics were grouped according to their sources (Sentinel 2A bands or vegetation indices) and regressed with field data to establish the strength of any relationships amongst the variables while using a non-parametric regression model (NPM). NPM allow for the modelling of densities and local polynomial regression on both continuous and categorical data, which do not necessarily follow any pre-defined distribution [55]. Hayfield and Racine [56] developed the *np* package in R used for this analysis. The procedure involves using the Akaike information criterion (AIC) to select the optimum bandwidths estimated from second-order Gaussian kernel densities. Local linear regressions are performed on the data while using the selected bandwidths to determine the fitting of the curve, and calculate the fitted, predicted, and error values. The *np* package has a multi-start function, which helps to avoid errors that occur in the presence of local minima.

The regression analysis was performed to investigate the possibility that the MSI detected changes in vegetation reflectance that are caused by oil pollution. Firstly, the spectral diversity of polluted transects should vary significantly with that of non-polluted transects. The rationale for this presumption is the documented effects of oil pollution on vegetation, which includes the decrease in plant productivity and loss of vulnerable plant species, thereby reducing the spectral diversity. Secondly, species diversity indices that were measured from the field were expected to linearly and positively regress with spectral diversity metrics (particularly on control transects) in line with the spectral variability hypothesis. Finally, selected spectral diversity metrics were predictors in models designed to test the SVH and estimate vascular plant species diversity in the study area. Metrics were selected based on the strength of their relationships with the diversity indices and the variance inflation factor (VIF). The VIF that was adopted to correct for collinearity in the explanatory variables was determined in R while using the *car* package. Only metrics with $VIF < 10$ were selected for the models. Prediction modelling involved a non-parametric multivariate regression (NPMR) analysis while using the *np* package in R. Two groups of spectral metrics namely, band-based (those derived from Sentinel 2A bands) and index-based (those derived from common vegetation indices) were used in models to estimate vascular plant species diversity. The dataset was randomly subdivided into training and test (validation data) while using a ratio of 7:3, respectively. Thus, the training data contained 150 observations and the test data contained 60 observations. Regression coefficients that were derived from the calibration process were applied to the test data for validation. The assessment of model performance was conducted by comparing the adjusted coefficient of determination values ($Adj.R^2$), root mean square error (RMSE), and predicted square error (PSE) of both band-based and index-based models.

The presence of spatial autocorrelation among the data points was acknowledged in this study and it was minimized by the selection of alternate segments on investigated transects. However, the spatial structure of the field data aggregated over transects and locations might lead to model over fit. Table 4 lists the models and parameters.

Table 4. Parameters of models used to estimate soil properties and other response variables from spectral diversity metrics computed using A. sentinel 2A bands 2 to 8A and B. common vegetation indices.

Model ID	Predictors	Response Variables
1A. Band-Based Metrics	Selected spectral metrics from Sentinel 2A bands 2 to 8A with $-0.2 < r > 0.2$	Species richness Species diversity Vegetation abundance
1B. Index-Based Metrics	Selected spectral metrics from vegetation indices with $-0.2 < r > 0.2$	Species richness Species diversity Vegetation abundance

3. Results

3.1. Total Petroleum Hydrocarbon (TPH) Distribution in Soil Samples

Figure 2 shows the average concentration of TPH in polluted and control transects across investigated locations. The highest concentration of TPH ($>99,000$ mg/kg) occurred in samples from the Amuruto spill epicentre. Overall mean and median TPH values for polluted transects were 12,692 and 3933 mg/kg, respectively, while for the control, the values were 40.53 and 36.50 mg/kg respectively. In Figure 2, the TPH concentration values are log-transformed and plotted to highlight the reference line, which signifies the average recommended Environmental Guidelines and Standards for the Petroleum Industry in Nigeria (EGASPIN) intervention value of different components of petroleum hydrocarbon in the soil. EGASPIN intervention values (mg/kg of soil) that were documented for different aromatic hydrocarbons that make up crude oil range from 1 for Benzene to 130 for Toluene (Department of Petroleum Resources, (DPR), 2002). The plot reveals that TPH concentration in polluted

transects are well above EGASPIN intervention values. Although lower TPH values were observed in control transects, these values were just within the borderline of intervention values and also above the target values of 0.05 mg/kg for the various petroleum hydrocarbon components.

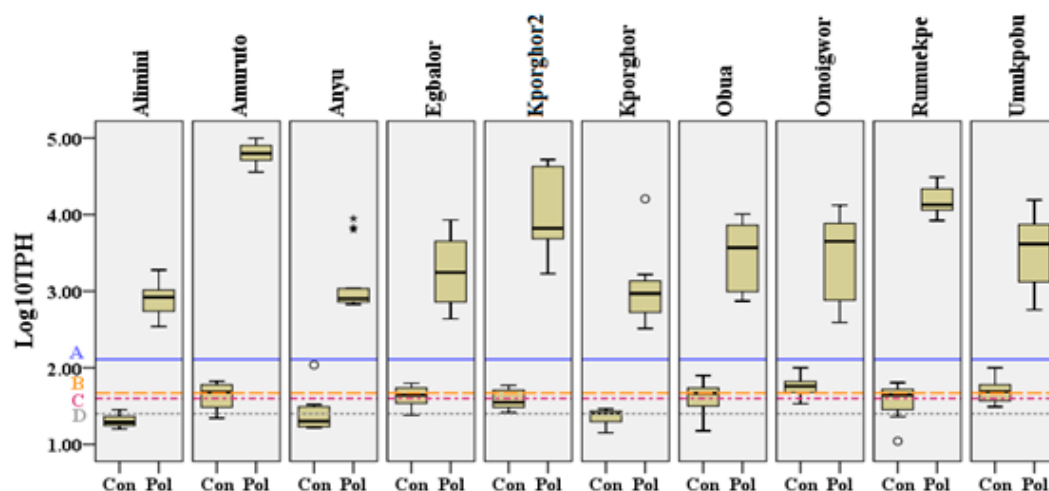


Figure 2. Total petroleum hydrocarbons levels in soil samples from polluted and control transects (N = 210). Bar charts were plotted using transformed (common log) values. The coloured reference lines illustrate the average Environmental Guidelines and Standards for the Petroleum Industry in Nigeria (EGASPIN) intervention value for A: Ethylbenzene = 50 mg/kg (1.669 on the log scale); B: Phenol = 40 mg/kg (1.6); C: Toluene = 130 mg/kg (2.11) and D: Xylene = 25 mg/kg (1.4). Total petroleum hydrocarbon (TPH) levels in polluted transects were well over the recommended intervention value, whereas levels in control 1 transects were borderline.

3.2. Species Composition, Richness and Diversity on Investigated Locations

163 plant species belonging to 52 families were recorded on transects. There were 37 families on polluted, and 52 on control (non-polluted) transects. In all of the locations, *Poaceae* was the most abundant family with 19 species. *Cyperaceae* followed with 13 species, then *Euphorbiaceae* and *Leguminosae* with ten species each. Other families with over five members were *Asteraceae*, eight; *Areaceae* and *Fabaceae*, seven each; *Malvaceae* and *Rubiaceae*, six each and *Sterculiaceae*, five. Species-wise, polluted transects had fewer species than control transects. The total number of different species (taxa) recorded on all the polluted transects was 93 and it was substantially lower than the 154 observed on all of the control transects. Generally, 11 species on polluted transects occurred in at least 30 segments out of 130 while on control transects, 15 species occurred in 30 segments out of 80. The most common species on the polluted transects was *Ageratum conyzoides* (Ageconh), and on control transects, it was *Manihot esculenta* (Manescs). At soil TPH levels greater than 50,000 mg/kg; only two species, namely *Costus afer* (Cosafes) and *Chloris pilosa* (Chlpilh), both annual plants occurred up to six times. Significant differences were apparent in vegetation characteristics that were measured from polluted and control transects. The species number (taxa) indicates the total number of unique species that occur on transects at the time of the investigation. Anyu had the highest taxa of 28 and 52 on polluted and control transects, respectively. The median taxa value for polluted and control segments were nine (n = 130) and 28 (n = 80), respectively.

The weighted average (WA) value of soil TPH (mg/kg) weighted by species abundance on segments was computed while using the *labdsv* package in R-studio to identify the most tolerant and most vulnerable species. The results in Table 5 shows the five most susceptible and five most tolerant species along with the WA scores. Interestingly, the most tolerant species were herbs, mainly *Perotis indica* (Perindh), which can tolerate over 67,000 mg/kg of TPH in the soil. The last letter in species code name indicates the life form of the species, s = shrub, h = herb, c = climber/creeper, t = tree). Other tolerant species include *Albizia adiantifolia* (Albadit), *Kyllinga erecta* (Kylereh), *Sida*

cordifolia (Sidcorh) and *Andropogon tectorum* (Andtech). The most susceptible species were *Terminalia catappa* (Tercatt), *Synedrella nodiflora* (Synnodh), *Oldenlandia corymbosa* (Oldcorh), *Albizia zygia* (Albzygt), and *Psychotria nigerica* (Psynigs).

Table 5. Weighted average scores of soil TPH concentration in investigated transects showing species most susceptible and most tolerant to oil pollution. The last letter in species code name indicates the life form of the species, s = shrub, h = herb, c = climber/creeper, t = tree).

Most Susceptible		Most Tolerant	
Species Code	WA Scores of TPH (mg/kg)	Species Code	WA Scores of TPH (mg/kg)
Tercatt	16.5	Perindh	67,940.25
Synnodh	19.2	Albadit	61,232.8
Oldcorh	25.9	Kylereh	22,048.08
Albzygt	26.26	Sidcorh	16,728.76
Psynigs	26.53	Andtech	14,470.86

Many indices were used to calculate the vascular plants species diversity of both the polluted and non-polluted transects across the entire study area. These included those for estimating species diversity, such as Shannon's (H) and Simpson's (D), those for estimating species richness, such as Menhinick's index (M) and Chao-1 (CH) indices. The analysis was performed while using PAST software using the abundance data of all inventoried species. Figure 3 shows the distribution of diversity indices values on polluted and control transects. From the plots, it is apparent that diversity values were higher on control transects than on polluted transect. Index values on polluted transects exhibited more variability than control transects, except in Chao-1, which showed a reverse with more variability in values that were obtained from control transects.

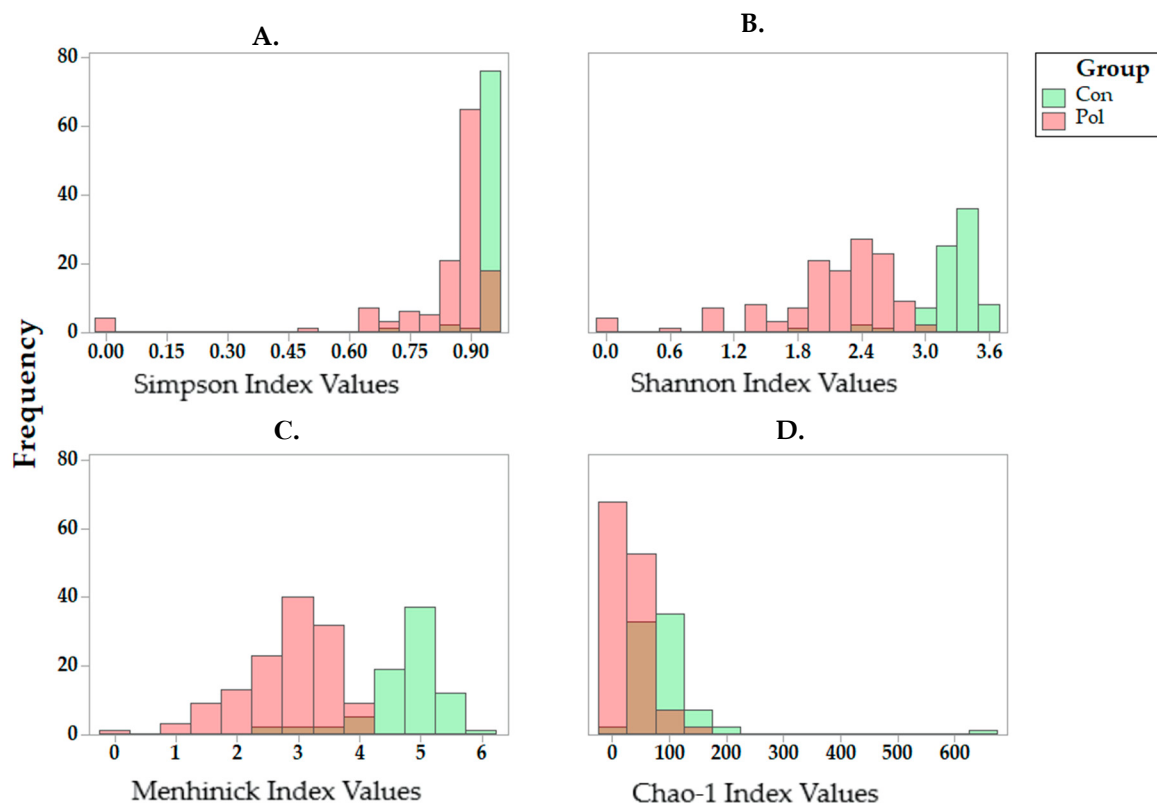


Figure 3. Histograms of values for (A). Simpson's, (B). Shannon's, (C). Menhinick's and (D). Chao-1 indices for polluted and control transects in investigated locations computed using the Paleontological Statistics Software Package, (PAST) software.

Generally, for all of the indices computed, the non-polluted transects exhibited higher richness and diversity values than polluted transects at all locations. Among the polluted transects, Menhinick's richness index ranged from 2.33 to 3.33, while the Shannon diversity values ranged from 0 to 2.93 computed for Rumuekpe. Among the non-polluted transects, species richness (Menhinick's Index) ranged from 3.26 to 5.42, and Shannon's index ranged from 1.87 to 3.6. Zero index values were mostly obtained on spill epicenters in Kporghor and Alimini, where fire incidence wholly removed the vegetation. The vegetation abundance on polluted transects appeared to be lower than on control transects. Median values were 1.27 plants per occurrence on control transects and 1.19 on polluted transects. The difference was significant when subjected to the Kruskal–Wallis test ($H = 76.06, p < 0.05$).

3.3. Relationship between Spectral Diversity Metrics and Species Richness/Diversity Indices

The median values of band-based spectral metrics from polluted transects were higher than those of the control transects. The exception were SH metrics and all of the metrics computed from band 8 (NIR) reflectance with higher median values in control transects (Table 6); however, the differences were not significant. The larger median values of polluted metrics were expected due to the increased RGB reflectance observed on polluted transects (supplementary material Figure S1). Similarly, the decreased NIR reflectance of polluted transects might have contributed to the reduced median values of polluted metrics in comparison to control metrics.

Table 6. Median values of some spectral metrics computed from Sentinel 2A bands showed higher values on polluted transects. The spectral heterogeneity (SH) metrics and all metrics computed from band 8 (NIR) reflectance (in bold) showed higher median values in control than in polluted transects although the differences were not significant.

Bands	Mean		Spectral Heterogeneity		1st Principal Component	
	Control	Polluted	Control	Polluted	Control	Polluted
2	1662.5	1678.5	0.03	−0.78	−0.21	−0.04
3	1516	1529	−0.17	−0.54	0.37	0.42
4	1357.5	1402.5	−0.11	−1.52	−0.27	0.19
5	1543.5	1618	−0.39	−2.22	−0.13	0.28
6	2217.5	2229.5	−0.32	−2.11	0.33	0.36
7	2455.5	2462	0.31	−0.43	0.36	0.38
8	2205	2199	−0.18	−3.58	0.38	0.37
8A	2595	2598	−1.58	−2.76	443	445

Correlation analysis showed linear relationships between band-based spectral diversity metrics and field measured vascular plant species diversity on both polluted transects and non-polluted transects. However, while strong and negative relationship prevailed on polluted transects; on control transects, they were mostly positive. This contradicted the expected strong and positive relationship on the control transects. The results of the Spearman's Rank Correlation analysis are illustrated in Figure 4. The plots were charted based on the metric derivation method, so as to identify the best performing metric. Each dot represents the r -value of a band metric versus the indicated species index on the X-axis. Dots in green are from control transects and those in red are from polluted transects. Labels on X-axis are a combination of index (Sm = Simpson, Sh = Shannon, Me = Menhinick's and Ch = Chao-1) and transect group (Con = Control, Pol = Polluted). The plots show that most of the spectral metrics correlated positively with indices on control transects and negatively with indices on polluted transects. The strongest positive relationships ($r > 0.3$) on control transects were from metrics based on the mean, median, and PC1 of bands, whereas the strongest inverse relationships ($r < -0.4$) that were observed on polluted transects were from metrics based on the spectral heterogeneity of

bands. Furthermore, the plots reveal that the best metrics were those computed from M, SD, PC1, and SH.

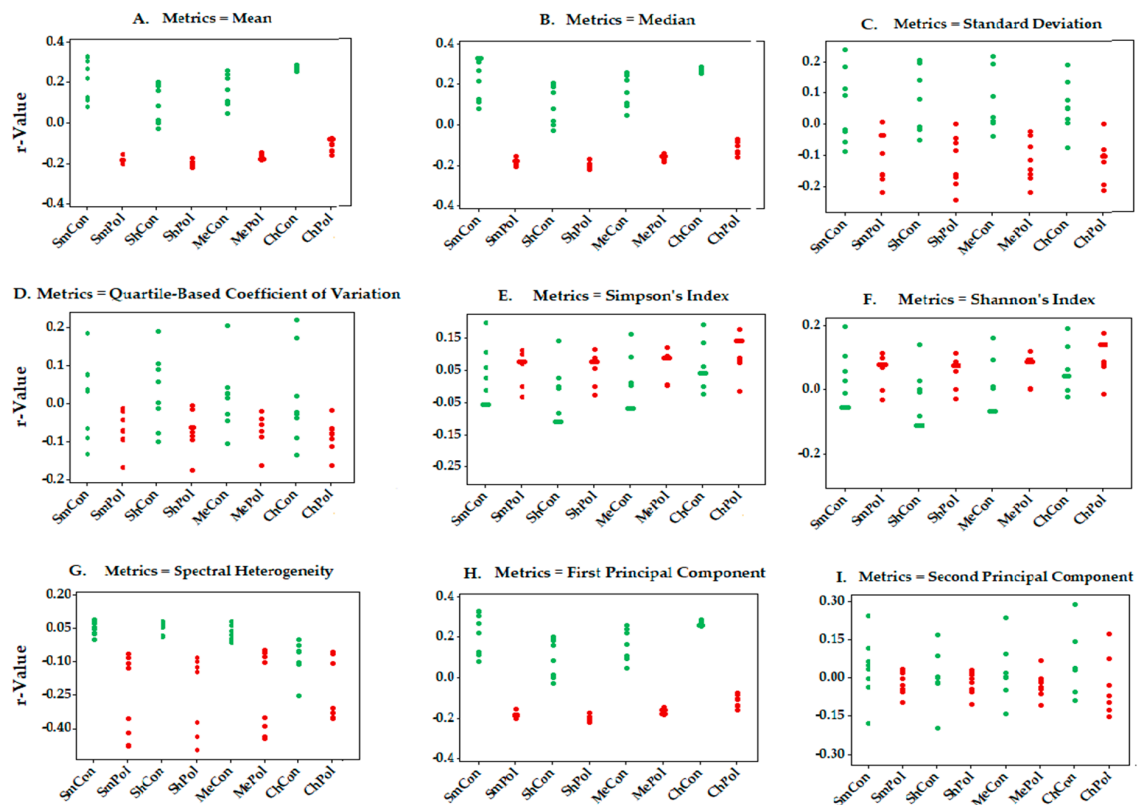


Figure 4. Dot plot of correlation coefficient (r) values of band-based spectral metrics and species diversity indices. The plot titles indicate the method of metric derivation (A). Mean, (B). median, (C). Standard Deviation, (D). Quartile-based Coefficient of Variation, (E). Simpson's index, (F). Shannon's index, (G). Spectral Heterogeneity, (H). First Principal Component and (I). Second Principal Component. Each dot represents the r -value of a band metric versus the indicated species index on the X-axis. Dots in green are from control transects while those in red are from polluted transects. The plots show that most spectral metrics correlated positively with indices on control transects and negatively with indices on polluted transects. Labels on X-axis are a combination of index (Sm = Simpson, Sh = Shannon, Me = Menhinick's and Ch = Chao-1) and transect group (Con = Control, Pol = Polluted).

The results of the NPM regression analysis using derivatives of bands and indices reveal strong relationships with the field measured diversity indices (R^2 values ranged from 0.19 to 0.98). The most robust relationships were found on control transects and involved metrics from both Sentinel 2A bands and vegetation indices. However this strong relationship is seen to depreciate across the study area when analysed with data from polluted and control transects. This suggests the sensitivity of the spectral metrics to the presence of soil TPH. Overall, the Simpson's index was the most sensitive variable with R^2 values that were greater than 0.5 for both metrics sets, except those derived from SIPI ($R^2 = 0.3$). The weakest relationships were between Chao-1 index and other spectral metrics (Figure 5).

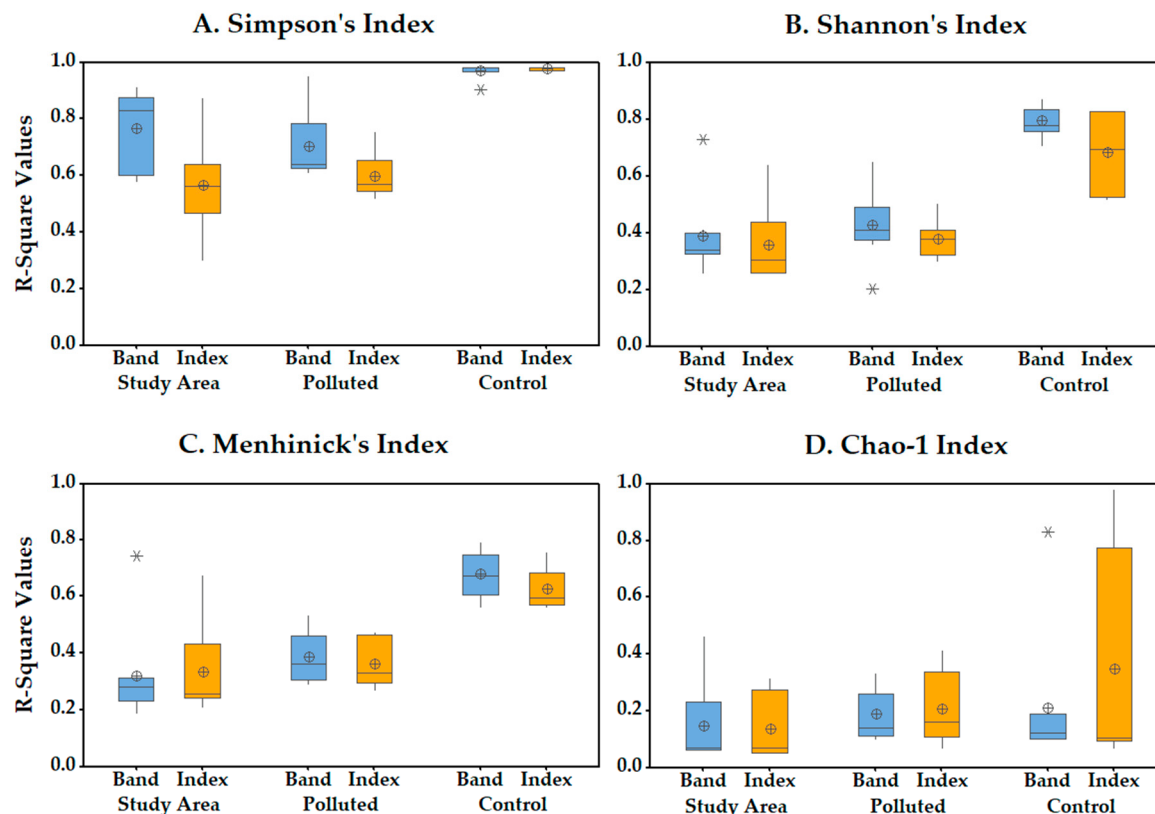


Figure 5. Boxplots of R-square values from regressing species diversity indices with spectral diversity metrics computed from Sentinel 2A bands and vegetation indices (A). Simpson's index, (B). Shannon's index, (C). Menhinick's index and (D). Chao-1 index. Band-based metrics are coloured blue and Index-based metrics are in orange. R-square values are compared for band and index-based metrics as well as among the study area, polluted transects and control transects. The plots show that stronger relationships were present on control transects than on polluted transects which in turn influenced a general weakening of this relationship across the study area. Overall, the Simpson's index was the most sensitive variable with R-square values greater than 0.5 for both metrics sets except those derived from SIPI (R-square = 0.3).

The most robust relationships occurred between Simpson's diversity index and metrics that were derived from the PC of bands ($R^2 = 0.91$) across the study area, Blue band ($R^2 = 0.89$) on polluted transects, and SIPI ($R^2 = 0.98$) on control transects. In contrast, the weakest relationships were observed between the Chao-1 richness index and spectral metrics from NDVI ($R^2 = 0.05$) across the study area, REP2 ($R^2 = 0.07$) on polluted transects, and CCI ($R^2 = 0.07$) on control transects. Spectral metrics that were derived from the PC of bands consistently exhibited a strong relationship with all of the field diversity indices with R^2 values well over 0.5, except for the Chao-1 index. Shannon's and Menhinick's indices both associated strongly with the transformed band metrics (PC of bands) and chlorophyll-based index (CCI) whereas Chao-1 richness index did not.

In terms of metrics performance, it appears that the band-based metrics were more sensitive to diversity indices than the index-based metrics across the study area. However, when separately analysed, the index-based metrics performed better on control transects, whereas band-based metrics were better on polluted transects. The boxplots in Figure 6 show that that the various spectral metrics performed better on control transects than on polluted transects. It appears that the metrics that were derived from stress indicating VIs (ARI and SIPI) were the most sensitive to the presence of soil TPH.

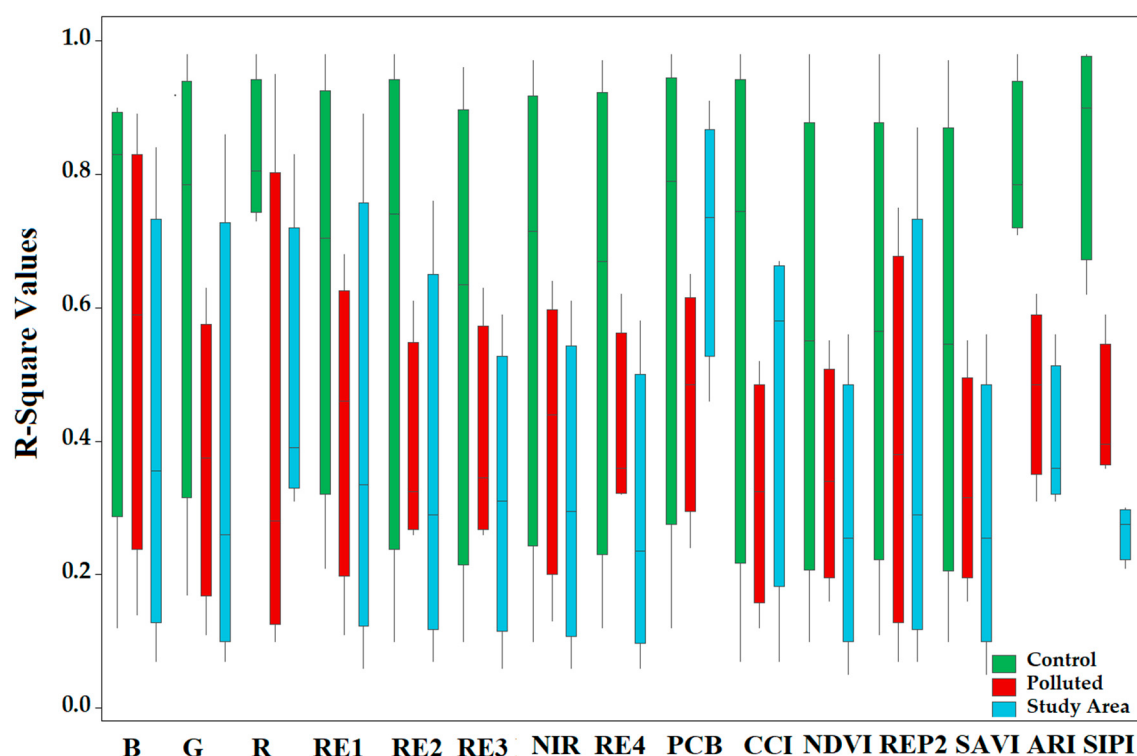


Figure 6. Boxplots of R^2 values illustrating the performance of metrics from individual bands and vegetation indices on polluted and control transects and across the entire study area. The plots show that the metrics were sensitive to the presence of TPH in the soil as strong relationships observed on control transects were weakened across the study area. The x-axis band labels are B = Blue, G = Green, R = Red, RE1 = Red Edge 1; RE2 = Red Edge 2; RE3 = Red Edge 3; NIR = Near Infrared; RE4 = Red Edge 4; PCB = Principal Components of Bands.

3.4. Estimating Species Diversity Using Spectral Diversity Metrics

The SVH was tested while using the spectral metrics that strongly correlated with the diversity measures ($r > \pm 0.2$) and a VIF < 10. The results in Table 7 show the R^2 values of training data and adjusted R^2 values test data, RMSE, and PSE of validation data sets for both models.

The results demonstrate that both sets of metrics performed well during model calibration with training data, but underperformed at model validation. For instance, among the diversity indices, Simpson's index is the most predictable with higher R^2 values (0.82) being obtained from calibrating models with band-based metrics and $R^2 = 0.63$ obtained from model calibration with index metrics. However, at validation, the adjusted R-square values were less than 0.5 for both sets of metrics, band-based metrics (Adj. $R^2 = 0.49$), and index-based metrics (Adj. $R^2 = 0.32$). For this index, band-based metrics were better estimators. Similar patterns were observed for Shannon's and Menhinick's indices with reduced Adj. R^2 values, although, index-based metrics outperformed band-based metrics as estimators of these indices. The least performing models were those that estimated the Chao-1 index. Despite high R^2 values at calibration, the models did not perform well at validation (results show no relationship among the variables). Graphical plots of residuals from model validation using test data are shown in the supplementary materials Figure S2. Model parameters are from the regression of spectral diversity metrics (A = band metrics and B = index metrics) on species diversity indices while using test data ($n = 60$).

Table 7. Performance summary of models estimating species richness and diversity using spectral diversity metrics computed from Sentinel 2A bands and vegetation indices. The combined dataset from polluted and control transects (N = 210) were employed in this analysis. The dataset was subdivided into training (n = 150) and test (n = 60) data. Explanatory variables (EV) for each model was a combination of all band-based or index-based metrics that showed a strong correlation ($r > \pm 0.2$) with the response variables.

Response Variable	Parameter	Band Metrics	Index Metrics
Simpson's	No of EV	9	11
	R ² (Train)	0.82	0.63
	Adj. R ² (Test)	0.49	0.32
	F	54.84 *	28.35 *
	RMSE	0.05	0.25
	PSE	0.002	0.06
Shannon's	No of EV	8	6
	R ² (Train)	0.36	0.41
	R ² (Test)	0.18	0.29
	F	12.81 *	22.82 *
	RMSE	0.68	0.63
	PSE	0.47	0.4
Menhinick's	No of EV	7	6
	R ² (Train)	0.29	0.57
	R ² (Test)	0.19	0.31
	F	14.26 *	27.12 *
	RMSE	1.11	1.03
	PSE	1.23	1.07
Log ₁₀ Chao-1	No of EV	5	10
	R ² (Train)	0.41	0.32
	R ² (Test)	0.02	0.07
	F	1.04ns	4.13 *
	RMSE	0.45	0.42
	PSE	0.2	0.18

*: significant; ns: not significant; bold: best performing model; shadow: unstable/least performing mode.

4. Discussion

The concept of the spectral variability hypothesis, as proposed by Palmer [1], provides ecologists with an essential tool for biodiversity monitoring and conservation in the Niger Delta region, without the need for labour intensive and time-consuming field work [20]. Understanding and modelling the relationship between spectral diversity metrics and local species richness or diversity measures provide decision makers with preliminary information regarding conservation priorities, particularly in cases where indicator species co-occur with rare or threatened species (cross-taxon surrogacy, [57]). Spectral variability hypothesis has been used several times to estimate species diversity and distribution in different landscapes and ecosystems. Hall et al. [21] reported that the area of the site might influence spectral variation of large sites; however, this phenomenon is not expected to occur within similar sized sample sites, as is the case in this study. The relationship between spectral variables and species diversity are usually weak, with R² values ranging from 0.2 to 0.6 [21]; however, the present study

reveals that a combination of various derivatives of spectral bands strongly associate with field data. Rocchini, Hernández-Stefanoni and He [57] stated that R^2 values of up to 0.5 could be considered to be valid to estimate species diversity from spectral variation, thus providing an integrated and efficient method for monitoring vascular plant species diversity at a regional and global scale.

4.1. Spectral Diversity Metrics for Estimating Species Richness and Diversity

Spectral diversity metrics correlated inversely with field measured vascular plant species diversity on polluted transects and positively on control transects. The coefficient of variation (CV) for metrics of polluted transects were larger than CV of metrics of control transects, and suggest that the polluted pixels were more diverse than control pixels. However, in the case of polluted transects, the spectral metrics do not necessarily depict vascular plant species diversity, but rather the heterogeneity of investigated transects. The reason for this is not far-fetched, since oil pollution accentuated habitat heterogeneity by creating patches of TPH-tolerant plants where TPH-susceptible species used to grow, thereby decreasing vegetation abundance and species diversity. Additionally, the weak positive correlation between the individual spectral metrics and species diversity indices of control transects might be ascribed to the greater homogeneity of segments on these transects in terms of species composition and distribution patterns. As the SVH relies on habitat heterogeneity, it follows that its application is limited in densely vegetated forests with little or no disturbance.

A combination of spectral diversity metrics successfully estimated the species richness and diversity of investigated locations with high R^2 values being obtained between the observed and predicted index values, despite the observed aberrations. The result was consistent with previous studies testing the SVH, such as Warren et al. [44]; Hall et al. [21]; Rocchini, Hernández-Stefanoni and He [57]; and, Schmidtlein and Fassnacht [58]. Index-based metrics outperformed band-based metrics in estimating Shannon's and Menhinick's indices, whereas band-based metrics did better at estimating the Simpson's diversity of investigated transects during validation. The Chao-1 richness index proved to be difficult estimating while using the computed metrics despite the model improvement measures undertaken.

4.2. Analysis of Multispectral (MS) Data for Detecting the Effects of Oil Pollution on Vegetation

Integrating remote sensing tools with field measurements yielded exciting results and highlighted its potential for biodiversity monitoring. Analyses of multispectral Sentinel 2A imagery using open-source software revealed the intricate connection between vegetation biochemical parameters and their spectral signatures. The study demonstrated the validity and applicability of the SVH in oil-polluted areas. On control transects, a positive linear relationship was expected because of the high species diversity, whereas, on polluted transects, an inverse relationship was expected due to low species diversity, but the results showed that the spectral diversity was equally high on polluted and control transects with the coefficient of variations (a measure of the dispersion of data points around the mean) for spectral metrics larger on polluted transects. This interesting result was attributed to the increased habitat heterogeneity following vegetation removal, waterlogging, and even burning of spill locations (habitat disturbance). Previous studies [5,6,59] show that these factors, measurable via remote sensing, influence species diversity. On control transects, the variations in the internal structures of different species, such as pigments and tissues that produce unique spectral signatures [20], controlled the relationship between spectral metrics and species diversity [20]. Similarly, the results also revealed that the SVH is sensitive to oil pollution effects on vegetation, given the change in the relationship that was observed between spectral diversity metrics and species diversity indices on polluted and control transects and across the study area. The regression results revealed that the strong relationship between variables on the control transects were weakened across the study area following the inclusion of metrics from polluted transects. Additionally, the results revealed an intricate relationship between metrics that were derived from stress indicating VIs (ARI and SIPI) and species diversity indices. As expected, the relationship was overwhelmingly negative on the control transects and mostly positive

on the polluted transects. Generally, stress indicating pigments, such as anthocyanin and carotenoid concentration in plants, increase when photosynthesis is impaired (a condition that can occur due to toxicity from oil pollution inhibiting chlorophyll production, Landi, Tattini, and Gould [60]). Hence the nature of the relationship between metrics derived from these VIs and diversity indices on polluted and control transects may be more of an indicator of oil pollution impact on vegetation health than on vascular plant species diversity. Previous studies testing the SVH achieved success with one set of metrics [21,44,57,58]. However, this study demonstrated that a combination of metrics derived from different statistical computations significantly strengthened the spectral diversity-species diversity relationship enough to estimate the Simpson's and Shannon's indices for the study area successfully.

4.3. Implications for Biodiversity Monitoring in the Niger Delta Region

Oil pollution in the Niger Delta region of Nigeria poses an existential threat to the people as well as the vast and diverse species of flora and fauna that inhabit the region. The results of the present study demonstrated the relevance of incorporating remote sensing technology in tackling critical environmental issues that are caused by oil pollution. The strength of relationships between spectral diversity metrics and field measured species diversity data can be exploited to develop solutions to environmental problems, such as halting species extinction through efficient monitoring and conservation policies. The peculiar condition of the Niger Delta region also demands alternative methods to traditional field survey practices that endanger lives. The SVH lends itself to several applications, including regular inspection of ecosystem services and biodiversity. Warren et al. [44] noted that plant species diversity is an essential indicator of ecosystem health, which can be monitored via the SVH.

The combination of spectral metrics showing strong relationships with species richness and diversity measures might be useful for mapping the species distribution of a given ecosystem. Schmidlein and Fassnacht successfully implemented a similar project [58] in mapping species occurrences in southern Germany while using multispectral data. Species distribution maps enhance conservation decisions, such as site prioritisation based on the structure and composition of plant communities revealed in the spectral variability of the maps [61].

Most importantly, the SVH is applicable in oil spill monitoring programmes to detect occurrences. The result of this study showed a clear distinction in the species composition of polluted and non-polluted transects, and this difference was apparent in the vegetation reflectance. Such definitive characterisation will enhance the monitoring of changes in polluted vegetation over time and space. Warren et al. [44] detected changes in species composition of a habitat subjected to different levels of disturbances.

5. Conclusions

Spectral metrics from Sentinel 2A validated the SVH and revealed the limitations to its application on polluted transects and on homogenous transects. Models that were based on these metrics successfully estimated vegetation parameters. However, the species-spectral diversity relationship was influenced by the presence of TPH in the soil, which challenged the SVH, as originally proposed in the literature. The results of the present study imply that vascular plants species diversity can be remotely estimated without going to the field to conduct vegetation surveys. This potential is very crucial in the Niger Delta region of Nigeria, where insecurity is a significant consideration in any field activity, thereby enhancing biodiversity monitoring over time and space.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/11/22/2662/s1>, Figure S1: Spectral profile of vegetation on polluted (red curves) and control (green curves) transects sampled from A. Rumuekpe, B. Umukpoku, C. Omoigwor and D. Alimini locations in Rivers State of Nigeria. Figure S2: Graphical plots of residuals from model validation using test data. Research Data available at <https://github.com/nmeticious/Research-Data>.

Author Contributions: Conceptualization, N.N.O.; Formal analysis, N.N.O.; Funding acquisition, N.N.O.; Methodology, N.N.O., H.B. and J.C.B.; Project administration, H.B.; Resources, H.B.; Supervision, H.B., J.C.B.; Validation, J.C.B.; Writing—original draft, N.N.O.; Writing—review & editing, N.N.O., H.B. and J.C.B.

Funding: This study was financed partly by the Tertiary Education Trust Fund (TETFUND), Michael Okpara University of Agriculture, Umudike, Nigeria, the National Center for Earth Observation (NCEO) and the Center for Landscape and Climate Research (CLCR), University of Leicester, United Kingdom.

Acknowledgments: The authors are grateful to the hardworking team of field assistants led by Victor Anih, the traditional rulers of various communities in Rivers state and the leadership of the youth groups for their support during the field campaign.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Palmer, M.W.; Earls, P.G.; Hoagland, B.W.; White, P.S.; Wohlgemuth, T. Quantitative tools for perfecting species lists. *Environmetrics* **2002**, *13*, 121–137. [[CrossRef](#)]
- Clevers, J.G.P.W.; De Jong, S.M.; Epema, G.F.; Van Der Meer, F.D.; Bakker, W.H.; Skidmore, A.K.; Scholte, K.H. Derivation of the red edge index using the MERIS standard band setting. *Int. J. Remote Sens.* **2002**, *23*, 3169–3184. [[CrossRef](#)]
- Sewelam, N.; Oshima, Y.; Mitsuda, N.; Ohme-Takagi, M. A step towards understanding plant responses to multiple environmental stresses: A genome-wide study. *Plant Cell Environ.* **2014**, *37*, 2024–2035. [[CrossRef](#)] [[PubMed](#)]
- Anjum, N.A. Plant acclimation to environmental stress: A critical appraisal. *Front. Plant Sci.* **2015**, *6*, 445. [[CrossRef](#)]
- Arellano, P.; Tansey, K.; Balzter, H.; Boyd, D.S. Detecting the effects of hydrocarbon pollution in the Amazon forest using hyperspectral satellite images. *Environ. Pollut.* **2015**, *205*, 225–239. [[CrossRef](#)] [[PubMed](#)]
- Pysek, P.; Pysek, A. Veränderungen der Vegetation durch experimentelle Erdgasbehandlung. *Weed Res.* **1989**, *29*, 193–204. [[CrossRef](#)]
- Scholten, M.C.T.; Leendertse, P.C. The impact of oil pollution on salt marsh vegetation. In *Ecological Responses to Environmental Stresses*; Rozema, J., Verkleij, J.A.C., Eds.; Springer: Dordrecht, The Netherlands, 1991; pp. 184–190.
- Beaubien, S.E.; Ciotoli, G.; Coombs, P.; Dictor, M.C.; Krüger, M.; Lombardi, S.; Pearce, J.M.; West, J.M. The impact of a naturally occurring CO₂ gas vent on the shallow ecosystem and soil chemistry of a Mediterranean pasture (Latera, Italy). *Int. J. Greenh. Gas Control* **2008**, *2*, 373–387. [[CrossRef](#)]
- Noomen, M.F.; van der Werff, H.M.A.; van der Meer, F.D. Spectral and spatial indicators of botanical changes caused by long-term hydrocarbon seepage. *Ecol. Inform.* **2012**, *8*, 55–64. [[CrossRef](#)]
- Zhang, C.G.; Leung, K.K.; Wong, Y.S.; Tam, N.F.Y. Germination, growth and physiological responses of mangrove plant (*Bruguiera gymnorhiza*) to lubricating oil pollution. *Environ. Exp. Bot.* **2007**, *60*, 127–136. [[CrossRef](#)]
- Rocchini, D.; Boyd, D.S.; Féret, J.; Foody, G.M.; He, K.S.; Lausch, A.; Nagendra, H.; Wegmann, M.; Pettorelli, N. Satellite remote sensing to monitor species diversity: Potential and pitfalls. *Remote Sens. Ecol. Conserv.* **2016**, *2*, 25–36. [[CrossRef](#)]
- Vihervaara, P.; Mononen, L.; Auvinen, A.; Virkkala, R.; Lu, Y.; Pippuri, I.; Packalen, P.; Valbuena, R.; Valkama, J. How to Integrate Remotely Sensed Data and Biodiversity for Ecosystem Assessments at Landscape Scale. *Landsc. Ecol.* **2014**, *30*, 501. [[CrossRef](#)]
- Mace, G.M.; Norris, K.; Fitter, A.H. Biodiversity and Ecosystem Services: A Multi-Layered Relationship. *Trends Ecol. Evol.* **2012**, *27*, 19–26. [[CrossRef](#)] [[PubMed](#)]
- Norris, K. Biodiversity in the Context of Ecosystem Services: The Applied Need for Systems Approaches. *Philos. Trans. R. Soc. Biol. Sci.* **2012**, *367*, 191–199. [[CrossRef](#)] [[PubMed](#)]
- Chapin, F.S.; Zavaleta, E.S.; Eviner, V.T.; Naylor, R.L.; Vitousek, P.M.; Reynolds, H.L.; Hooper, D.U.; Lavorel, S.; Osvaldo, E.S.; Hobbie, S.E.; et al. Consequences of Changing Biodiversity. *Nature* **2000**, *405*, 234–242. [[CrossRef](#)]
- Waide, R.B.; Willig, M.R.; Steiner, C.F.; Mittelbach, G.; Gough, L.; Dodson, S.I.; Juday, G.P.; Parmenter, R. The Relationship between Productivity and Species Richness. *Annu. Rev. Ecol. Syst.* **1999**, *30*, 257–300. [[CrossRef](#)]

17. Pereira, H.M.; Cooper, D.H. Towards the Global Monitoring of Biodiversity Change. *Trends Ecol. Evol.* **2006**, *21*, 123–129. [\[CrossRef\]](#)
18. Jetz, W.; McGeoch, M.A.; Guralnick, R.; Ferrier, S.; Beck, J.; Costello, M.J.; Fernandez, M.; Geller, G.N.; Keil, P.; Merow, C.; et al. Eren Turak Essential biodiversity variables for mapping and monitoring species populations. *Nat. Ecol. Evol.* **2019**, *3*, 539–551. [\[CrossRef\]](#)
19. Onyia, N.; Balzter, H.; Berrio, J. Normalized Difference Vegetation Vigour Index: A New Remote Sensing Approach to Biodiversity Monitoring in Oil Polluted Regions. *Remote Sens.* **2018**, *10*, 897. [\[CrossRef\]](#)
20. Heumann, B.W.; Hackett, R.A.; Monfils, A.K. Testing the spectral diversity hypothesis using spectroscopy data in a simulated wetland community. *Ecol. Inform.* **2015**, *25*, 29–34. [\[CrossRef\]](#)
21. Hall, K.; Reitalu, T.; Sykes, M.T.; Prentice, H.C. Spectral heterogeneity of QuickBird satellite data is related to fine-scale plant species spatial turnover in semi-natural grasslands. *Appl. Veg. Sci.* **2012**, *15*, 145–157. [\[CrossRef\]](#)
22. Saatchi, S.; Buermann, W.; ter Steege, H.; Mori, S.; Smith, T.B. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sens. Environ.* **2008**, *112*, 2000–2017. [\[CrossRef\]](#)
23. NEDECO (Netherlands Engineering Consultants). *Rivers Studies and Recommendations on Improvement of Niger and Benue*; North-Holland: Amsterdam, The Netherlands, 1959.
24. Omo-Irabor, O.O.; Olobaniyi, S.B.; Akunna, J.; Venus, V.; Maina, J.M.; Paradzayi, C. Mangrove Vulnerability Modelling in Parts of Western Niger Delta, Nigeria Using Satellite Images, GIS Techniques and Spatial Multi-Criteria Analysis (SMCA). *Environ. Monit. Assess.* **2011**, *178*, 39–51. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Elenwo, E.; Akankali, J. Impact of Climate Change on Aquatic Fauna of Economic Importance in Niger Delta, Nigeria. *Atmos. Clim. Sci.* **2014**, *4*, 710–720. [\[CrossRef\]](#)
26. *Environmental Assessment of Ogoniland Report*; United Nations Environmental Programme: Nairobi, Kenya, 2011; ISBN 978-92-807-3130-9.
27. Nigerian Oil Spill Monitor. Available online: <https://oilspillmonitor.ng/> (accessed on 5 October 2015).
28. Adishi, E.; Hunga, M.O. Oil Theft, Illegal Bunkering and Pipeline Vandalism: It's Impact on Nigeria Economy, 2015–2016. *Int. J. Econ. Bus. Manag.* **2017**, *3*, 47–65.
29. Cummings, J.; Smith, D. *The Line-Intercept Method: A Tool for Introductory Plant Ecology Laboratories*; Association for Biology Laboratory Education (ABLE), Clemson University: Clemson, SC, USA, 2000; pp. 234–246.
30. Kercher, S.M.; Frieswyk, C.B.; Zedler, J.B. Effects of sampling teams and estimation methods on the assessment of plant cover. *J. Veg. Sci.* **2003**, *14*, 899–906. [\[CrossRef\]](#)
31. Tanee, F.B.G.; Albert, E. Reconnaissance Assessment of Long-Term Effects of Crude Oil Spill on Soil Chemical Properties and Plant Composition at Kwawa, Ogoni, Nigeria. *J. Environ. Sci. Technol.* **2015**, *8*, 320–329. [\[CrossRef\]](#)
32. Ubom, R.M. Ethnobotany and Biodiversity Conservation in the Niger Delta, Nigeria. *Int. J. Bot.* **2010**, *6*, 310–322.
33. Agbagwa, I.O.; Ekeke, C. Structure and Phytodiversity of Freshwater Swamp Forest in Oil-rich Bonny, Rivers State, Nigeria. *Res. J. For.* **2011**, *5*, 66–77. [\[CrossRef\]](#)
34. Rocchini, D.; Balkenhol, N.; Carter, G.A.; Foody, G.M.; Gillespie, T.W.; He, K.S.; Kark, S.; Levin, N.; Lucas, K.; Luoto, M.; et al. Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecol. Inform.* **2010**, *5*, 318–329. [\[CrossRef\]](#)
35. Small, C. The Landsat ETM+ spectral mixing space. *Remote Sens. Environ.* **2004**, *93*, 1–17. [\[CrossRef\]](#)
36. Rocchini, D. Effects of spatial and spectral resolution in estimating ecosystem α -diversity by satellite imagery. *Remote Sens. Environ.* **2007**, *111*, 423–434. [\[CrossRef\]](#)
37. Turner, W.; Spector, S.; Gardiner, N.; Fladeland, M.; Sterling, E.; Steininger, M. Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* **2003**, *18*, 306–314. [\[CrossRef\]](#)
38. Chen, W.; Henebry, G.M. Change of spatial information under rescaling: A case study using multi-resolution image series. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 592–597. [\[CrossRef\]](#)
39. Parviainen, M.; Luoto, M.; Heikkinen, R.K. The role of local and landscape level measures of greenness in modelling boreal plant species richness. *Ecol. Model.* **2009**, *220*, 2690–2701. [\[CrossRef\]](#)
40. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel Algorithms for Remote Estimation of Vegetation Fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [\[CrossRef\]](#)

41. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating Chlorophyll Content from Hyperspectral Vegetation Indices: Modeling and Validation. *Agric. Meteorol.* **2008**, *148*, 1230–1241. [[CrossRef](#)]
42. Mielke, S.P.; Kiang, N.Y.; Blankenship, R.E.; Gunner, M.R.; Mauzerall, D. Efficiency of Photosynthesis in a Chl D-Utilizing Cyanobacterium is Comparable to Or Higher than That in Chl A-Utilizing Oxygenic Species. *Biochim. Biophys. Acta (BBA) Bioenerg.* **2011**, *1807*, 1231–1236. [[CrossRef](#)]
43. Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
44. Warren, S.D.; Alt, M.; Olson, K.D.; Irl, S.D.H.; Steinbauer, M.J.; Jentsch, A. The Relationship between the Spectral Diversity of Satellite Imagery, Habitat Heterogeneity, and Plant Species Richness. *Ecol. Inform.* **2014**, *24*, 160–168. [[CrossRef](#)]
45. Ringnér, M. What is principal component analysis? *Nat. Biotechnol.* **2008**, *26*, 303–304. [[CrossRef](#)]
46. Pu, R.; Gong, P.; Yu, Q. Comparative Analysis of EO-1 ALI and Hyperion, and Landsat ETM+ Data for Mapping Forest Crown Closure and Leaf Area Index. *Sensors* **2008**, *8*, 3744. [[CrossRef](#)] [[PubMed](#)]
47. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
48. Rouse, J.; Haas, R.; Schell, J.; Deering, D. *Monitoring Vegetation Systems in the Great Plains with ERTS; Third Earth Resources Technology Satellite-1 Symposium—Volume I: Technical Presentations*; NASA Goddard Space Flight Center: Washington, DC, USA, 1974; pp. 309–317.
49. le Maire, G.; François, C.; Dufrêne, E. Towards Universal Broad Leaf Chlorophyll Indices Using PROSPECT Simulated Database and Hyperspectral Reflectance Measurements. *Remote Sens. Environ.* **2004**, *89*, 1–28. [[CrossRef](#)]
50. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
51. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [[CrossRef](#)] [[PubMed](#)]
52. Peñuelas, J.; Filella, I. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci.* **1998**, *3*, 151–156. [[CrossRef](#)]
53. Wallis, W.A.; Kruskal, W.H. Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583.
54. Dunn, O.J. Multiple Comparisons Using Rank Sums. *Technometrics* **1964**, *6*, 241–252. [[CrossRef](#)]
55. Albek, E. Estimation of Point and Diffuse Contaminant Loads to Streams by Non-Parametric Regression Analysis of Monitoring Data. *Water Air Soil Pollut.* **2003**, *147*, 229–243. [[CrossRef](#)]
56. Hayfield, T.; Racine, J.S. Nonparametric Econometrics: The np Package. *J. Stat. Softw.* **2008**, *27*, 1–32. [[CrossRef](#)]
57. Rocchini, D.; Hernández-Stefanoni, J.L.; He, K.S. Advancing species diversity estimate by remotely sensed proxies: A conceptual review. *Ecol. Inform.* **2015**, *25*, 22–28. [[CrossRef](#)]
58. Schmidtlein, S.; Fassnacht, F.E. The spectral variability hypothesis does not hold across landscapes. *Remote Sens. Environ.* **2017**, *192*, 114–125. [[CrossRef](#)]
59. Douda, J.; Doudová-Kochánková, J.; Boublík, K. Alena Drašnarová Plant species coexistence at local scale in temperate swamp forest: Test of habitat heterogeneity hypothesis. *Oecologia* **2012**, *169*, 523–534. [[CrossRef](#)] [[PubMed](#)]
60. Landi, M.; Tattini, M.; Gould, K.S. Multiple functional roles of anthocyanins in plant-environment interactions. *Environ. Exp. Bot.* **2015**, *119*, 4–17. [[CrossRef](#)]
61. Rocchini, D.; Chiarucci, A.; Loiselle, S.A. Testing the spectral variation hypothesis by using satellite multispectral images. *Acta Oecol.* **2004**, *26*, 117–120. [[CrossRef](#)]

