ISSN No:-2456-2165

Comparing Ordinary Least Square Regression and GWR for Modelling NDVI-Precipitation Relationships over Crop/Grassland Ecosystem in Northwestern Nigeria

Sa'ad Ibrahim[1,*], Babangida Malik[1], Umar Mohammed Lawal [2], Ibrahim Sa'adu [3] , Abdullahi Mohammad [1]
[1]Department of Geography, Adamu Augie College of Education, PMB 1012, Argungu, Kebbi State, Nigeria
[2]Department of Biology, Adamu Augie College of Education, PMB 1012, Argungu, Kebbi State, Nigeria:
[3]Department of Geography, Usmanu Danfodiyo University, PMB 2346, Sokoto, Nigeria

Abstract:- The crop/grassland of savannah is a complex ecosystem in which the relation between vegetation productivity and precipitation is uncertain due to high interannual climate variability and anthropogenic activities. This posed a serious threat to biodiversity, food security and socioeconomic development. In view of this, previous studies have emphasized on identifying the effective modelling approaches for quantifying these relationships, mostly by comparing the ordinary least square (OLS) and geographically weighted regression (GWR) models. Although, the conventional regression failed to successfully model these relationships, most previous research who compare the two techniques for studying the influence of precipitation on vegetation only used normalized vegetation difference index (NDVI) metrics for various locations without referring to specific vegetation type. In this study, we investigated the relationships between the NDVI metrics acquired from a 15-year Moderate resolution imaging spectroradiometer (MODIS) time series data and mean total annual precipitation generated from the inverse distance weighted interpolation technique using the ground observations data (15 years) of 42 weather stations in Nigeria. The study compared OLS and GWR modelling approaches in a crop/grassland dominated savannah. OLS did not find any significant relationship between the NDVI metrics and mean total annual precipitation. In contrast, the GWR modelling shows that the relationship exists. The rainfed crops (R^2 = 0.66), mosaic croplands/vegetation (R^2 = 0.65) and mosaic vegetation/croplands (R^2 =58) were found to respond more strongly to mean total annual precipitation using GWR. The GWR found the highest R^2 values of 0.66 and 0.97 for the individual observations and global estimate respectively. The rainfed crops, mosaic croplands/vegetation and mosaic vegetation/croplands showed the largest variability, and were much more sensitive to variability in the precipitation than other vegetation types.

I. INTRODUCTION

In recent years, there has been an increasing interest for understanding the distribution and condition of plant function types as a function of environmental variables such as rainfall. Over the years, climate is being considered as one of the most important factor influencing the growth and distribution of plant species. A study using dynamic vegetation models (DVGMs) which grow plants with reference to physiological principles using climate and soil as input, predict that vast wooded grasslands in Africa, South America and other smaller areas of grassy biomes have climate potential to form forest (Bond and Keeley, 2005). The study offers some important insights into regional assessments of vegetation and potential factors influencing its distribution.

This view was further supported by Sankaran et al., (2005) who investigated the determinants of woody cover in over 854 savannah sites in African. The research highlights the influence of resource (such as water, nutrients, fire and herbivore etc.) availability and the distribution of plant species within savannah ecosystem. The analysis demonstrated that the tree cover is not simply associated to resource abundance, that for all sites with <650mm mean annul precipitation (MAP), tree cover is constrained linearly with moisture availability whereas for the sites having >650mm MAP canopy closure is possible. Disturbances such as fire can prevent canopy closure. Therefore, climate, soil nutrients, fire are some of the essential components that monitor savannah dynamics. The spatial pattern of rainfall being one of the most important component in the climate system, is related to NDVI metrics (e.g. maximum, monthly NDVI, mean annual NDV etc.) and is said to be largely responsible for the strong variation in vegetation canopy phenology (Chudumayo 2001). The relationship between these components is however a function of time scale, topography, vegetation composition and structure (Chamaille-Jammes et al., 2006). Thus, the NDVI-precipitation relationships is site-specific. NDVI-precipitation relationship in Eastern and Southern African savannas was studied by Chamaille-Jammes and Fritz (2009). Result indicates that only semi-arid savannah has

ISSN No:-2456-2165

significant precipitation –NDVI relationships. In other parts, the scenario was that maximum NDVI was opposite to rainfall conditions at MAP gradient. These results provided further support on the earlier hypothesis on African savannah that ecosystems sensitivity decreases with increasing MAP (Chamaille-Jammes and Fritz, 2009). This means that crop/grassland ecosystems would respond more strongly to precipitation changes than a homogenous ecosystem such as forest.

The impact of precipitation extremes, for example, has been and is still being used to assess vegetation vulnerability (Liu, Liu, & Yin, 2013; Propastin, Fotso, & Kappas, 2010). Liu et al. (Liu, et al., 2013) studied the global patterns of NDVI while trying to find out the impact of precipitation extremes on PFTs in order to assess the vulnerability of ecosystems. Results suggest that vegetation in temperate broadleaf forest and temperate grassland is highly prone to extreme events under more severe precipitation extremes. The study suggests that more attention should be paid to precipitation-induced vegetation changes than to temperature-induced events in those regions. Propastin et al. (Propastin, et al., 2010) investigated the vulnerability of PFTs over Africa to El-Niño Southern Oscillation (ENSO) events for the period 1982-2006 using the moving-window statistical correlation analysis technique. The sensitivity of PFTs to ENSO events were assessed using NDVI data from Advanced Very High-Resolution Radiometer (AVHRR) and ENSO indices. Results indicated that the vulnerability of vegetation in Africa depends to a large extent on the distribution of PFTs. Grasslands, closed shrublands and woodlands exhibit the largest share of areas with moderate and high vulnerability. Evergreen broadleaved forest and deciduous broadleaved forest areas were significantly less affected by ENSO than the other land cover types during the study period. This means that the grassland ecosystems of savannah where agriculture is mainly practiced by the dwellers is more vulnerable. This posed significant threat to food security and socio-economic development.

Several attempts were made in the past to examine the relationships between NDVI and precipitation across varying gradient. However, a number of these studies (Plessis 1999: Balaghi et al. 2004: Jammes & Fritz 2009) have relied on the Ordinary Least Square (OLS) Regression which give a global approach to measurements and therefore do not take account of the spatial non-stationarity of the data being modelled. Conversely, geographically weighted regression (GWR) is a technique that advanced regression analysis by accounting for spatial non-stationarity of the data. GWR is provides a more local rather than global analysis (Foody 2006). The application of GWR is however not new in the field of remote sensing. But most previous studies that applied this technique for studying the influence of precipitation on vegetation used NDVI metrics for various locations without referring to specific vegetation type. In this study, five structural vegetation types in a crop/grassland ecosystem were considered. The objectives of this study were to investigate: (1) the differences between the OLS and GWR in modelling the NDVI-precipitation relationship in different vegetation types. (2) the response of the vegetation type to precipitation based on the NDVI metrics extracted from the Moderate Resolution Imaging Spectroradiometer (MODIS) time series data : and (3) to assess which of the NDVI metric is the main phenology indicator that is influencing the relationships in these structural vegetation types.

II. STUDY AREA

The area used in this study is located in Kebbi state, (12° - 13°N and 4°- 5° E), which is the north-western part of Nigeria. The climate of the area is tropical continental type; with two clear marked seasons, the dry and wet seasons resulting from two contrasting air masses, the tropical continental and tropical maritime originating from the Sahara Desert and the Atlantic Ocean respectively. The wet season lasts May-October with mean annual rainfall of 800 mm. Agriculture is the mainstay of farmers growing crops such as millet, rice, cassava, sugar cane, maize, guinea corn etc. We extracted 30 locations of five major vegetation types (Open to closed grassland, mosaic vegetation/croplands, rainfed croplands, mosaic croplands/vegetation, and Sparse vegetation) in the study area (Table 1), from the GlobCover Land Cover Map, 2006 which provided by the European Space Agency. GlobCover 2006, is part of the European Space Agency GlobCover Projects. This land cover map was produced for the entire world is processed in an automated chain from MERIS with the view to provide detailed land information and also serves as an enabling platform for time series analysis. Since inception, the products of this project (including the GlobCover 2006) have been receiving positive feedback and interesting comments from a large partnerships and reputable international institutions (e.g. Food Agricultural Organisation, United Nation Environment Programmes, and International Geosphere Biosphere Programme etc.).

ISSN No:-2456-2165

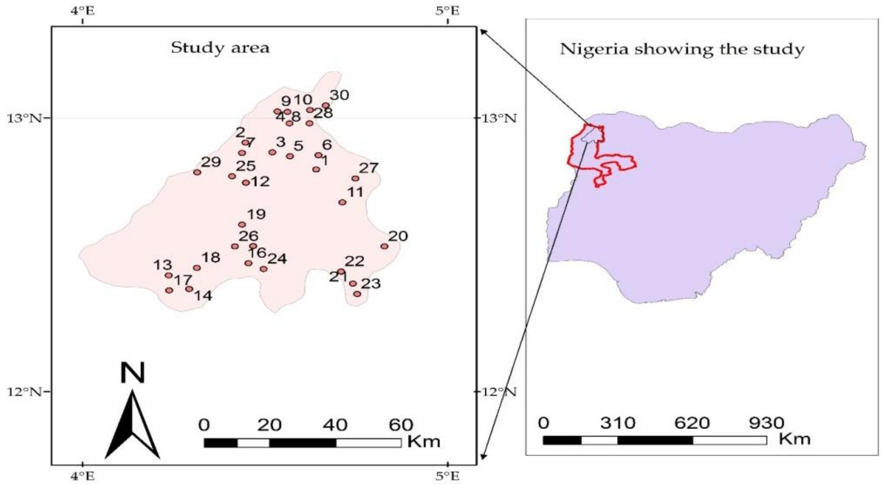

Fig 1:- Map of the study area showing the locations investigated.

A. Remote sensing and meteorological data

➢ *MODIS NDVI time series data*

MODIS time series NDVI data was acquired from January 2001 to December 2015 (DAAC, 2019; Didan, 2015) (via https://modis.ornl.gov/cgibin/MODIS/global/subset.pl). It is MOD13Q1, a gridded level 3 product provided at 250 m spatial resolution every 16 days produced from atmospherically corrected bi-directional surface reflectance factors (BRFs) and masked for water, clouds, and cloud shadows. For this study, NDVI metrics were calculated from the MODIS time series data. The metrics include the mean annual, mean seasonal and maximum NDVI of a 15-year MODIS time series data. Mean NDVI is the mean for all observations within the time series. The mean seasonal NDVI, which is the mean of the rainy season images for the time series data. While, the maximum NDVI, which is the mean of the images with the maximum values computed from the time series data.

S/No	Longitude (°)	Latitude (°)	PREC.	M. NDVI	M. SEAS. NDVI	MAX. NDVI	Vegetation type
1	4.64	12.812	853.483	0.263	0.423	0.5575	Mosaic Vegetation/Croplands
2	4.446	12.91	866.335	0.247	0.412	0.5833	Rainfed croplands
3	4.52	12.875	864.185	0.317	0.503	0.6245	Mosaic Vegetation/Croplands
4	4.567	12.981	844.255	0.334	0.573	0.6793	Rainfed croplands
5	4.568	12.86	859.218	0.224	0.32	0.4212	Closed to open grassland
6	4.646	12.864	843.475	0.367	0.495	0.6344	Rainfed croplands
7	4.437	12.872	878.192	0.213	0.351	0.4368	Mosaic Vegetation/Croplands
8	4.561	13.022	841.171	0.297	0.563	0.6947	Rainfed croplands
9	4.534	13.024	846.59	0.309	0.574	0.6515	Rainfed croplands
10	4.623	13.029	829.875	0.253	0.444	0.6228	Mosaic Vegetation/Croplands
11	4.712	12.692	860.334	0.259	0.502	0.6147	Closed to open grassland
12	4.448	12.764	886.12	0.265	0.414	0.5009	Mosaic Vegetation/Croplands
13	4.236	12.425	949.164	0.4	0.54	0.6435	Mosaic Croplands/Vegetation
14	4.292	12.375	951.528	0.285	0.508	0.5728	Mosaic Vegetation/Croplands
15	4.467	12.532	920.071	0.268	0.462	0.5317	Mosaic Vegetation/Croplands
16	4.454	12.469	925.263	0.309	0.463	0.605	Mosaic Vegetation/Croplands
17	4.237	12.37	953.421	0.424	0.552	0.5931	Rainfed croplands
18	4.313	12.452	940.438	0.283	0.441	0.5489	Mosaic Vegetation/Croplands
19	4.437	12.611	908.384	0.249	0.439	0.5138	Closed to open grassland
20	4.827	12.531	881.78	0.266	0.48	0.5408	Mosaic Croplands/Vegetation
21	4.741	12.395	920.36	0.278	0.481	0.6313	Mosaic Vegetation/Croplands
22	4.708	12.439	909.885	0.291	0.333	0.5222	Mosaic Vegetation/Croplands
23	4.752	12.357	924.52	0.27	0.423	0.5274	Mosaic Croplands/Vegetation
24	4.496	12.448	927.792	0.295	0.44	0.5712	Mosaic Vegetation/Croplands
25	4.41	12.787	886.064	0.236	0.378	0.4691	Mosaic Vegetation/Croplands
26	4.418	12.531	923.047	0.25	0.369	0.4197	Closed to open grassland
27	4.747	12.779	837.399	0.227	0.38	0.5108	Sparse vegetation
28	4.622	12.981	833.082	0.245	0.43	0.5396	Closed to open grassland
29	4.314	12.801	898.033	0.244	0.415	0.5373	Mosaic Vegetation/Croplands
30	4.666	13.046	821.069	0.273	0.464	0.6203	Closed to open grassland

Table 1

➢ *Precipitation data*

Precipitation data was acquired from the Nigerian Meteorological Agency (NIMET) of Nigeria. Daily precipitation of 42 weather stations of Nigeria from 2001 to 2015 were collected. Mean annual precipitation were then calculated and georeferenced. The georeferenced data were interpolated using inverse distance weighted. Inverse distance weighted was reported as one of the most accurate interpolation techniques (Chen and Liu, 2012). The 30 locations were chosen based on the major land cover types of the study area and their corresponding precipitation data were extracted. Table 1 shows mean total annual precipitation value for each location.

B. Methods

➢ *OLS*

This regression technique is aspatial. It does not consider the spatial locations of the individual observation to be modelled. In this study, the mean total annual total precipitation was taken as the explanatory variable while mean, seasonal or maximum NDVI metric as the response variables.

➢ *GWR*

The relationships between the NDVI metrics and precipitation were assessed using the GWR model. GWR model relationships at varying spatial scale. This means that unlike, the OLS, GWR is spatial, and therefore takes account of non-stationarity of the data being examined (Foody 2003). The coefficient of determination for each observation in thirty (30) locations was assessed using each NDVI metric and their corresponding mean total annual precipitation. In addition, the global R^2 value for each NDVI metric was also recorded.

III. RESULTS

A. Relationship between NDVI metrics and mean annual precipitation using OLS regression

The OLS regression analysis revealed no significant relationship between NDVI metrics and the mean annual precipitation (15 years of ground observations) precipitation (Figure 2). Specifically, the mean annual NDVI versus the precipitation shows a coefficient of determination R^2 =0.09, p =0.09. The relationships for the mean seasonal and maximum seasonal NDVI with the mean total annual precipitation indicated an R^2 =0.0001, p =0.99 and R^2 =0.04, p =0.28 respectively.

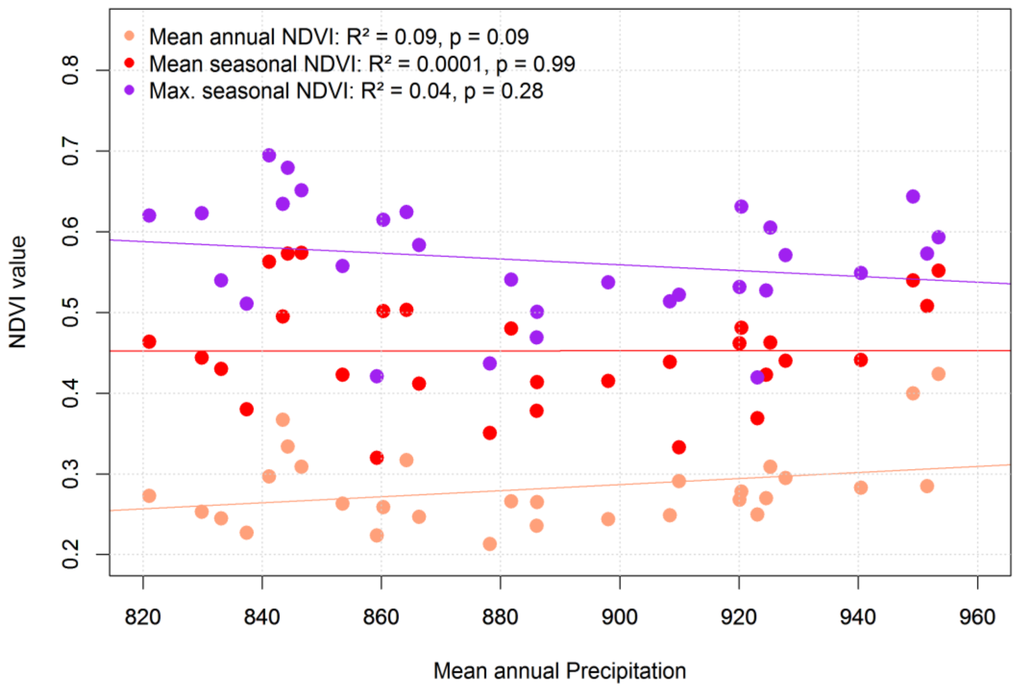

Fig 2:- Relationship between NDVI metrics and mean total annual precipitation using OLS regression

B. Relationship between NDVI metrics and mean annual precipitation using GWR regression

The relationships between the NDVI metrics and mean total annual precipitation using GWR is shown on Figure 3a/c. Since GWR is spatial, each individual observation has its own R-squared value depending upon the nature of the two interacting variables. The mean NDVI shows an R^2 of 0.45 and 0.38 for the mosaic croplands/vegetation and grassland respectively. The mean seasonal NDVI, as expected, had the highest R^2 (0.66) (Figure 3b). The four land cover classes with the highest R^2 for the seasonal mean NDVI metric include the rainfed crops (R^2 = 0.66), Mosaic croplands/vegetation (R^2 = 0.65) and mosaic vegetation/croplands (R^2 =58), While the maximum NDVI also had rainfed crops with the highest R^2 of 0.39 (Figure 3c). In the same NDVI metric, mosaic vegetation/croplands croplands/ mosaic vegetation had R^2 of 0.35 each (Figure 3c). Grassland (closed to open grassland) had an R^2 of more than 0.38 which is the highest compared to other NDVI metrics (Figure 3a/c).

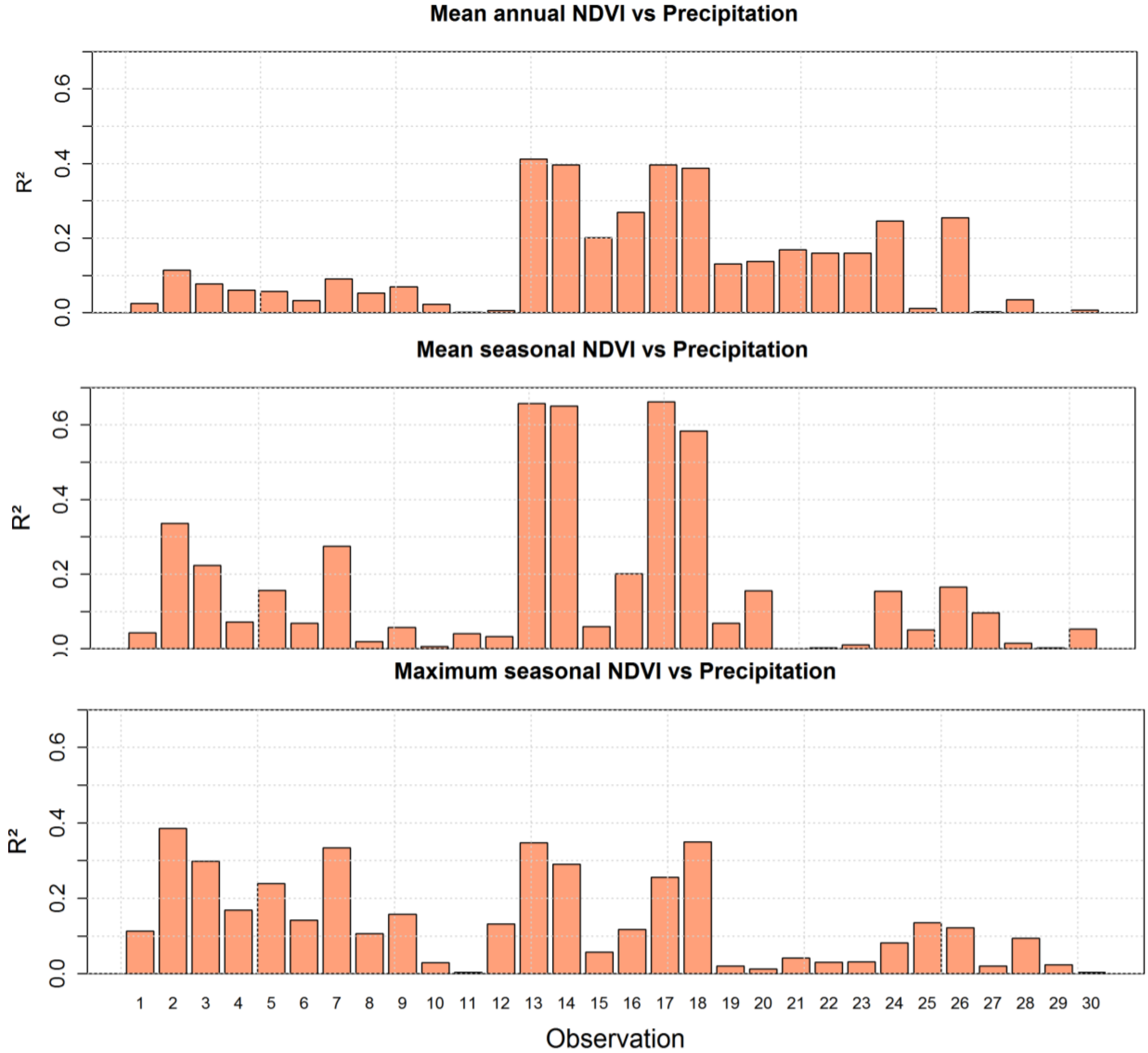

Fig 3: Relationship between NDVI metrics and mean total annual precipitation using GWR regression

C. Comparison of NDVI-precipitation relationships between OLS and GWR

The regression analyses results for the two models (OLS and GWR) were further compared using the global R^2. GWR despite being local, the model also provides overall R^2 values for the observations. The mean annual NDVI metric shows an R^2 of 0.09 for the OLS while in GWR the R^2 = 0.88. Using the mean seasonal NDVI value, OLS had least R^2 while GWR provides the best estimate (R^2 =0.95). For the maximum NDVI metric an R^2 of 0.04 and 0.94 were reported for OLS and GWR respectively.

NDVI metric	OLS		GWR	
	R^2	R^2 adjusted	R^2	R^2 adjusted
Mean Annual NDVI	0.09	0.07	0.91	0.88
Mean Seasonal NDVI	0.000001	-0.04	0.97	0.95
Maximum NDVI	0.04	0.007	0.96	0.94

Table 2: NDVI-precipitation relationships using OLS and GWR

IV. DISCUSSION

This study investigated the relationships between the NDVI metrics and their corresponding mean total annual precipitation in parts of Kebbi state, Nigeria. The results from the two modelling approaches, the Ordinary Least Square (OLS) and geographically weighted regression (GWR) were compared. The study found out that the GWR is more promising in quantifying the NDVI-precipitation relationships than the OLS.

Results from the OLS modelling approach indicated no significant relationship between the NDVI metrics of different vegetation types (dominated by crop/mosaic vegetation and grassland) and mean total annual precipitation. In contrast, the GWR modelling shows that strong relationship exists. One, the GWR modelling approach shows that the extent to which the variability of the NDVI metrics of a given vegetation type can be explained by the mean total annual precipitation. The rainfed crops (R^2 = 0.66), mosaic croplands/vegetation (R^2 = 0.65) and mosaic vegetation/croplands (R^2 =58) were found to respond more strongly to mean total annual precipitation. These findings are consistent with those of Foody (2003) who investigate NDVI- rainfall relationship, explained that OLS is a poor local descriptor of relationship (R^2 =0.67) compared to GWR (R^2 =0.96). Propastin & Kappas (2008) also found GWR more accurate than the OLS, in which OLS recorded an R^2 value of 0.75 and 0.63 for the individual class and whole land cover types respectively. While GWR increases the accuracy level of this relation to R^2 values to 0.97. This is simply because global models are limited in explaining the relationship between set of variables due its universal approach (Fotheringham, Brunsdon and Charlton, 2003).

Secondly, GWR revealed that the seasonal mean NDVI metric is the main phenology indicator that influence the relationship with these structural vegetation types. The highest R^2 values are is 0.66 and 0.97 for the individual observations and global estimate respectively. Previous researchers have also examined the differences between the OLS and GWR in evaluating the non-stationarity relationships between the mean annual precipitations and irrigated and rainfed Maize and Soybeans yields (Sharma et al., 2019). The OLS shows that rainfed crops has an R^2 =0.67 whereas R^2 =0.23 and R^2 =0.17 were found for irrigated maize and soybean, respectively. The performance of the GWR technique in predicting the same yields for irrigated and rainfed maize and soybean was significantly better than the performance of the OLS model. The R^2 for the rainfed maize and soybean is in the range of 0.77 to 0.80 while for irrigated crops, is 0.42. GWR is considers local condition of the individual observation. Recently, Lembrechts, Nijs and Lenoir (2019) have emphasized on incorporating microclimate into species distribution models.

This study shows that GWR modelling approach outperformed the OLS in estimating the relationships between the NDVI metrics and mean total annual precipitation in a crop/grass-dominated savannah. The quantitative analysis of modelling these relationships between the NDVI metrics from satellite time-series data and mean total annual precipitation enables a better understanding of the response of crop/grass-dominated savannah to weather and climate variability.

V. CONCLUSION

In this study, we investigated the relationships between the NDVI metrics acquired from a 15-year MODIS time series data and mean total annual precipitation generated from the inverse distance weighted interpolation of ground observations data (15 years) of 42 weather stations in Nigeria. The study compared OLS and GWR modelling approaches in a crop/grassland dominated savannah. OLS modelling approach indicated no relationship between the NDVI metrics of different vegetation types (dominated by crop/mosaic vegetation and grassland) and mean total annual precipitation. In contrast to OLS, the GWR modelling shows that the relationship exists. The GWR modelling approach shows the extent to which the variability of the NDVI metrics of the vegetation type can be explained by the mean total annual precipitation. The rainfed crops (R^2 = 0.66), mosaic croplands/vegetation (R^2 = 0.65) and mosaic vegetation/croplands (R^2 =58) were found to respond more strongly to mean total annual precipitation. GWR revealed that the seasonal mean NDVI metric is the main phenology indicator that influence the relationships in these structural vegetation types. The GWR found the highest R^2 values of 0.66 and 0.97 for the individual observations and global estimate respectively. The rainfed crops, mosaic croplands/vegetation and mosaic vegetation/croplands showed the largest variability, and were much more sensitive to variability in the precipitation than other vegetation types.

FUNDING

This research was funded by the Tertiary Education Trust Fund (TETFUND), Nigeria, through the Institution Based Research Fund (IBRF).

ACKNOWLEDGMENTS

This research was supported by TETFUND. We acknowledge the R Core team for making various library packages available for the data analyses. We also acknowledge the European Space Agency (ESA) for providing the GlobCover land cover map. Special thanks to National Aeronautics and Space Administration (NASA) for providing the MODIS data.

REFERENCES

[1]. Bond, W.J. and Keeley, J.E., 2005. Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends in ecology & evolution*, **20**(7), pp. 387-394.

[2]. Chamaille-jammes, S., Fritz, H. and Murindagomo, F., 2006. Spatial patterns of the NDVI-rainfall relationship at the seasonal and interannual time scales in an African savanna. *International Journal of Remote Sensing*, 27(23), pp. 5185-5200.

[3]. Chamaille-jammes, S. and Fritz, H., 2009. Precipitation-NDVI relationships in eastern and southern African savannas vary along a precipitation gradient. *International Journal of Remote Sensing*, 30(13), pp. 3409-3422.

[4]. Chen, F.W, Liu C.W. 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy and Water Environment*, **10**: 209– 222. DOI 10.1007/s10333-012-0319-1.

ISSN No:-2456-2165

[5]. Chen, Q., 2007. Airborne LiDAR data processing and information extraction. *Photogrammetric Engineering and Remote Sensing*, 73(2), pp. 109.

[6]. Chidumayo, E.N., 2001. Climate and Phenology of Savanna Vegetation in Southern Africa. *Journal of Vegetation Science*, **12**(3), pp. 347-354

[7]. DAAC, O. (2019). MODIS and VIIRS Land Products Global Subsetting and Visualization Tool. ORNL DAAC Retrieved January 21, 2019, from https://doi.org/10.3334/ORNLDAAC/1379

[8]. Didan, K. (2015). . MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC (Publication no. 10.5067/modis/mod13q1.006).

[9]. Du Plessis, W.P., 1999. Linear regression relationships between NDVI, vegetation and rainfall in Etosha National Park, Namibia. *Journal of Arid Environments*, **42**(4), pp. 235-260.

[10]. Fotheringham AS, Brunsdon C, Charlton M 2003. Geographically weighted regression: the analysis of spatially varying relationship. Wiley, New York

[11]. Lembrechts, J.J., Nijs, I. and Lenoir, J. (2019), Incorporating microclimate into species distribution models. *Ecography*, 42: 1267-1279. doi:10.1111/ecog.03947

[12]. Liu, G., Liu, H., & Yin, Y. (2013). Global patterns of NDVI-indicated vegetation extremes and their sensitivity to climate extremes. *Environmental Research Letters*, *8*(2), 025009. doi: 10.1088/1748-9326/8/2/025009

[13]. Propastin, P., Fotso, L., & Kappas, M. (2010). Assessment of vegetation vulnerability to ENSO warm events over Africa. *International Journal of Applied Earth Observation and Geoinformation*, 12, 83-89. doi: https://doi.org/10.1016/j.jag.2009.10.007

[14]. Sharma, V., Irmak, A., Kabenge, I., & Irmak, S. (2011). Application of GIS and Geographically Weighted Regression to Evaluate the Spatial Non-Stationarity Relationships Between Precipitation vs. Irrigated and Rainfed Maize and Soybean Yields. *Transactions of the ASABE*, 54(3), 953-972. doi: https://doi.org/10.13031/2013.41227

[15]. Sankaran, M., Hanan, N., Scholes, R. et al. Determinants of woody cover in African savannas. *Nature* 438, 846–849 (2005) doi:10.1038/nature04070.