

## Application of Ordinary Kriging in Mapping Soil Organic Carbon in Chad using SoilGrids data

Kella Douzouné<sup>1</sup>, Joseph Oloukoi<sup>2</sup>, Emmanuel Ehnong Gongnet<sup>3</sup>, and Tranquillin Sèdjro Antoine Affossogbe<sup>4</sup>

<sup>1</sup> African Regional Institute for Geospatial Information Science and Technology (AFRIGIST), Obafemi Awolowo University Campus, Ile-Ife, Nigeria, kella@afrigist.org

<sup>2</sup> African Regional Institute for Geospatial Information Science and Technology (AFRIGIST), Obafemi Awolowo University Campus, Ile-Ife, Nigeria, oloukoi@afrigist.org

<sup>3</sup> Institut Tchadien de Recherche Agronomique pour le Développement, Tchad

<sup>4</sup> Laboratoire de Biomathématiques et d'Estimations Forestières, Faculty of Agronomic Sciences, University of Abomey-Calavi Campus (LABEF/FSA/UAC), Building CBIG (2nd floor), Abomey-Calavi, O4 BP 1525 Cotonou, Bénin

DOI: <https://dx.doi.org/10.4314/sajg.v13i2.11>

### ABSTRACT

*The quantification of the pattern and spatial distribution of soil organic carbon (SOC) is fundamental to understanding many ecosystem processes. This study aimed to apply ordinary kriging (OK) to model the spatial distribution of SOC in Chad. A total of 995 sampling locations from the region were used to extract soil organic carbon from three raster layers. Those raster layers represented the SOC of 0-5 cm, 5-15 cm, and 15-30 cm of soil horizon and were downloaded from the SoilGrids website. The mean value of the soil carbon derived from the three horizons was used as 0-30cm horizon data and analysed using R-4.1.3 version software and ArcGIS 10.5. Different variogram models were first examined on the variogram cloud, and, based on RMSE, MSE, and MAE criteria, the best fit was selected. The results indicated that the Gaussian model is the best fit to the data, with 27.84, -3.35, and 20.95 obtained, respectively, for RMSE, MAE, and ME. The short-range spatial dependence of SOC was strong, with a nugget close to zero. The spatial dependency of the data was medium, with a nugget-to-sill ratio of 0.36. The southern portion of the country has a higher concentration of SOC than the northern portion. It can be concluded that the generated map could serve as a proxy for SOC in the region where evidence of spatial structure and quantitative estimates of uncertainty are reported. Therefore, the maps produced can be used for many applications, including soil sampling optimization.*

**Key Words:** Gaussian semi-variogram, spatial distribution, soil organic carbon, spatial autocorrelation

## **1. INTRODUCTION**

Agriculture is the main source of financial income for almost 80% of the poor population. It remains a major sector in the economies of West African countries (Guèye, 2003). For good physical and mental health, people need to consume healthy and rich food. However, the variability of soil fertility caused by climate change effects alters nutritional quality, yield, and crop productivity (Razakarison, 2013), thus increasing the risk of food insecurity (N'guessan & Bernard, 2012). Also, there is strong demographic pressure that requires increased agricultural yields. To solve this problem, many farmers turn to cultivation practices involving the excessive use of chemical fertilizers. The knowledge of the specific location where fertilisers are needed on a given farm is important for soil fertility management. According to Jennings et al. (2009), the understanding of the spatial distribution of soil fertility parameters for soil productivity is very important in agriculture. The need to boost crop productivity and reduce environmental risk is a source of great interest to scientists (Du et al., 2016). Soil chemical and physical parameters such as nitrogen (N), organic carbon (SOC), phosphorus (P), and potassium (K) are important macronutrients for plant growth (Li et al. 2016; Tripler et al. 2006). As demonstrated by Marklein and Houlton (2012), nitrogen inputs have been shown to accelerate phosphorus cycling rates across various ecosystems, thus underscoring the complex interdependencies among these essential nutrients. Soil chemical and physical parameters such as nitrogen (N), organic carbon (SOC), phosphorus (P), and potassium (K) are important macronutrients for plant growth (Li et al., 2016; Tripler et al., 2006). In addition, they are the main elements of soil which help to describe soil fertility across a given space. According to Wei et al. (2008), soil organic carbon (SOC) is mostly used to characterize soil quality in ecological modeling and environmental prediction for precision agriculture. According to Gouri et al. (2016), knowledge of the SOC spatial pattern offers tools for evaluating soil fertility and sustainable agriculture. It is also applied in natural resource management (Zhang et al., 2012, Liu et al., 2014). However, many problems affect the spatial variability evaluation of SOC. One of the problems is its higher spatial variability in a soil unit (Ceri et al., 2000). In addition, the analysis of the soil sampling data is very costly. Despite these aspects, adequate information about the spatio-temporal behavior of soil parameters is needed (Chabala et al., 2017). This necessitates regional studies that allow for the refining of global estimates, which are obtained from aggregated regional data (Clothing et al., 2013).

Several studies have applied geostatistical methods (Kriging) to evaluate soil organic carbon concentrations in various regions of the world. In the Iran region, Mahmoud Zadeh (2020) evaluated the spatial prediction of soil organic carbon using machine learning techniques in western Iran; Chabala et al. (2017) used kriging techniques to map soil organic carbon in Zambia; Owusu et al. (2020) predicted the spatial distribution of SOC in Ghana using a regression-kriging model; Zhang (2010) combined an explanatory variable with the kriging method to predict SOC spatial distribution in China; and Sreenivas et al. (2016) mapped the

concentration of SOC in India and Akpa et al. (2020) in Nigeria. Some of these studies have shown that geostatistical methods such as ordinary kriging and regression kriging, enhanced by environmental variables, improve the accuracy of SOC predictions [Boubehziz et al., 2020; Addise et al., 2022]. These techniques are crucial, especially in regions prone to erosion and intensive agricultural practices. In Sub-Saharan Africa, geochemical properties and climatic variables have been identified as key predictors of SOC variation, thus providing crucial insights for land management strategies (von Fromm et al., 2021). It is globally observed that the spatial variation of SOC and other soil fertility parameters is under-evaluated in some African countries. This is widely due to the unavailability of freely accessible soil data. Since the introduction in 2016 of the *SoilGrids* soil data platforms, the application of soil data to understand the dynamics of each soil parameter has become more attractive. The objectives of this study are to apply ordinary kriging to model the distribution of SOC in Chad based on SoilGrids data.

Our study provides vital information for agricultural producers who require a precise understanding of the spatial distribution of SOC to optimize the application of soil amendments, reduce costs, and increase agricultural productivity by adopting these geostatistical methods. By providing accurate SOC maps, we not only help improve soil management, but also plan agricultural interventions more sustainably, enabling farmers to better respond to the challenges of climate change and soil degradation.

## **2. MATERIALS AND METHODS**

### **2.1. Study area**

Chad occupies an area of about 1,284,000 km<sup>2</sup> in Africa. It extends between latitudes 7° and 24°N and longitudes 13° and 24°E. It is bordered by Libya in the north, Sudan in the east, Niger, Nigeria, and Cameroon in the west, and by the Central African Republic in the south. The country is divided into three zones. The northern portion is a desert, while the middle portion accommodates an arid Sahelian belt. A fertile zone where the majority of the population lives (Sudanian savanna) is located in the south, where the annual rainfall ranges between 600 and 1,300 mm. The primary sector here contributes to the livelihoods of almost 75% of the population living in rural areas. Agriculture is mainly dominated by cotton, sorghum, millet, maize, peanuts, rice, potatoes, gum Arabic, and tapioca (Ba, 2017), while the animal livestock component comprises cattle, sheep, goats, and camels.



Figure 1: Study area.

## 2.2. Data source and analysis

On a regular grid size of 100 m<sup>2</sup>, 1,000 soil sample locations were projected on the map of Chad, thus ensuring comprehensive coverage and statistical representation across various soil and land use types. The geographic coordinates of each location were extracted and reported. To determine SOC values for soil horizons of 0-30 cm, three different soil organic carbon horizon rasters (0-5 cm, 5-15 cm, and 15-30 cm) were accessed and downloaded from the SoilGrids website (<https://soilgrids.org/>). Because the soil horizons of 0–30 cm were not available on the SoilGrids website, the mean values of the three SOC horizons were computed from the three distinct layers at each point to effectively represent the SOC content. This methodological choice is supported by the fact that most plant roots in agricultural settings

extend up to 30 cm deep in the soil, making this depth particularly relevant for SOC analysis. In addition to that, owing to the unavailability of soil organic carbon values at some locations, 994 sampling points were finally retained for analysis (Fig.1). All procedures were performed using R-4.1.3 for statistical analysis and ArcGIS 10.5 for spatial mapping and analysis. After this preliminary screening, summary statistics (mean, standard deviation, skewness, kurtosis, minimum and maximum) were generated to provide a basic understanding of the characteristics of the data. The spatial distribution of SOC across the study area was conducted using ArcGIS 10.5 (Jarnevich et al., 2018).

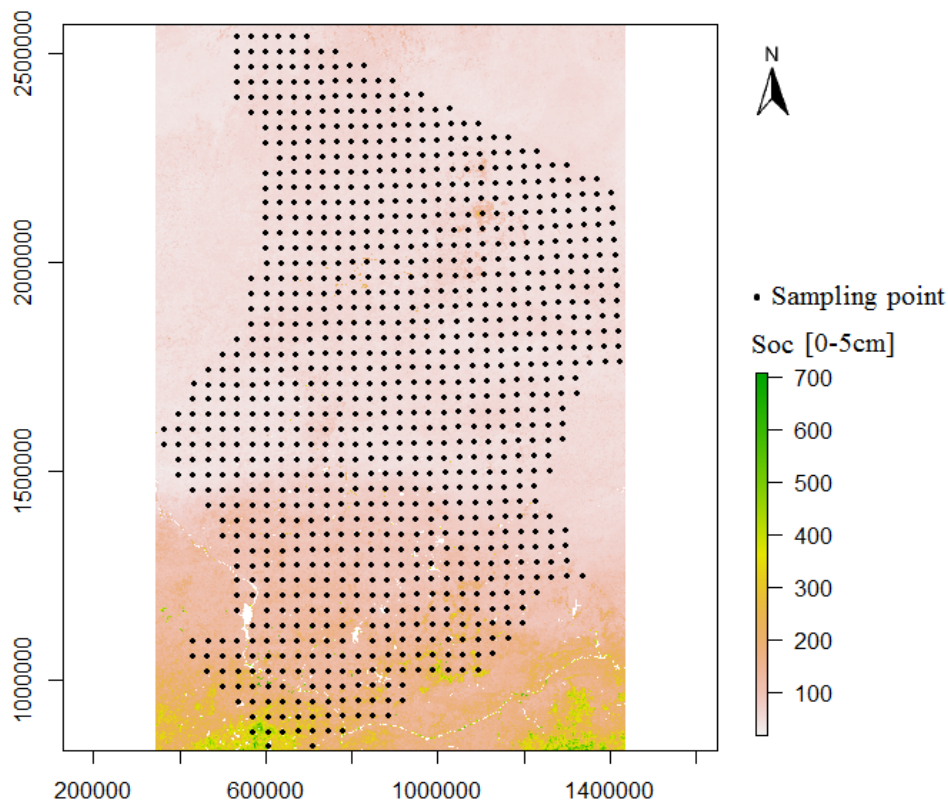


Fig. 2: Location of sampling points recorded from the map of Chad and projected on the SOC horizon, 0-5cm.

### **2.3. Ordinary kriging for the spatial distribution of SOC**

Ordinary kriging (OK) was employed to evaluate the spatial variability of SOC in Chad. This geostatistical interpolation method uses information related to a given phenomenon at a certain location to predict the value of a given phenomenon (in this case, SOC) at an unsampled location (Ryu et al., 2002). Its selection was based on its suitability for managing sparse data in large datasets, like those typical of environmental and soil studies (Nagle, 2010). The kriging process involves several steps: first, constructing a semi-variogram to model spatial dependence; second, using the semi-variogram model to estimate the kriging weights; and

finally applying these weights to predict SOC at unsampled points. Bhunia et al. (2018) provides a detailed comparison of GIS-based interpolation methods, further validating the choice of kriging for accurate spatial predictions. In addition to that, it is the most used kriging technique for soil property prediction (Yimit et al., 2011). The kriging process involves several steps: first, constructing a semi-variogram to model spatial dependence; second, using the semi-variogram model to estimate the kriging weights; and finally, applying these weights to predict SOC at unsampled points. The accuracy of these predictions was quantified using cross-validation techniques, in which predictions at known points were compared to actual measured values.

Let us denote by  $Z'(x_0)$ , the predicted soil property at unsampled location  $x_0$ , and by  $\sigma_k^2(x_0)$ , the error variance of the prediction computed at each sampled location using measured values,  $Z(x_i)$  ( $i = 1; 2; n$ ), by means of the following equation:

$$Z'(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (1)$$

Where:

$Z'(x_0)$ : Estimated value of the soil property at location  $x_0$ ,

$\lambda_i$ : Kriging weight for the sampled location  $i$ ,

$Z(x_i)$ : Measured value of the soil property at the sampled location  $i$ ,

$n$ : Number of sampled locations.

$$\sigma_k^2(x_0) = \mu + \sum_{i=1}^n \lambda_i \gamma(x_0 + x_i) \quad (2)$$

Where:

$\mu$  is the lag range constant,

$\lambda_i$  is the kriging weight, and

$\gamma(x_0 + x_i)$  is the semi-variogram value corresponding to the distance between  $x_0$  and  $x_i$ .

At this level, the semi-variograms represent the basic tool which is used to evaluate the spatial distribution of SOC. According to Nielsen and Wendroth (2003), it is expressed as follows:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (3)$$

Where

$\gamma(h)$  is the semi-variance,

$h$  is the lag distance,

$Z$  is the parameter of the soil property,

$N(h)$  is the number of pairs of locations separated by a lag distance,  $h$ ,

$Z(x_i)$  and  $Z(x_i + h)$  are values of  $Z$  at positions  $x_i$  and  $x_i + h$  (Wang and Shao, 2013).

### 2.3.1. Model validation

To evaluate the most appropriate SOC data, various semi-variogram models, including exponential, Gaussian, spherical, penta-spherical, and cubic models, are tested. The performance of the models was compared using statistical measures such as the mean absolute error (MAE), the mean error (ME), and the root mean square error (RMSE), expressed as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [Z(x_i) - Z'(x_i)]^2} \quad (4)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |Z(x_i) - Z'(x_i)| \quad (5)$$

$$ME = \frac{1}{N} \sum_{i=1}^N Z(x_i) - Z'(x_i) \quad (6)$$

With  $\Delta$  denoting the range (the difference between the maximum and minimum observed data),  $Z(x_i)$  the measured value of SOC, and  $Z'(x_i)$  the predicted value.

## 3. RESULTS AND DISCUSSION

### 3.1. Summary statistics

The statistical analysis of soil organic carbon (SOC) data from SoilGrids data revealed that the mean SOC across Chad was 56.73 g/kg with a standard deviation of 33.49 g/kg. The distribution showed significant skewness (2.5 g/kg) and kurtosis (8.18 g/kg), indicating that SOC values in Chad are highly variable and not normally distributed. This variability is typical in large-scale environmental datasets and emphasizes the need for robust geostatistical methods to accurately model SOC distribution.

Table1: Summary statistic of soil organic carbon in Chad

Statistics	Observed values	Transformed values (logarithmic)
Mean	56.73	3.92
Standard error	1.06	0.014
Standard deviation	33.49	0.44
Kurtosis	8.18	0.71
Skewness	2.5	1.21
Minimum	27	3.29
Maximum	289	5.67
Sample size	994	994

### 3.2. The variography of the SOC and the kriging interpolation surface

During the identification of the best-fit variogram model to the data, many models are plotted on the data variogram cloud (fig. 3a) and the appreciation parameters (RMSE, MAE and ME), reported in Table 2. The SOC variogram analysis proved to be critical in understanding the spatial dependency and in structuring the interpolation process. We evaluated several variogram models, including Gaussian, exponential, and spherical, to determine the best fit for our data. The Gaussian model was selected on the basis of its lowest root mean square error (RMSE) of 27.84, its mean error (ME) of -3.35, and its mean absolute error (MAE) of 20.95, thus suggesting that it could provide the most reliable predictions with minimal bias (Boubehziz *et al.*, 2020) (Fig. 3b and Table 2). This choice was justified by the ability of the model to handle the medium-range spatial dependence observed in our SOC data which is characterized by a nugget-to-sill ratio of 0.36, suggesting that while there is some degree of spatial randomness, much of the variability in SOC can be explained by spatial autocorrelation. This finding supports the use of ordinary kriging for SOC prediction in regions with similar spatial characteristics (Addise *et al.*, 2022). According to Gambardella *et al.* (1994), the spatial distribution of specific data in each area could be judged to be high if the nugget-to-sill ratio is less than 0.25, medium if it is between 0.25 and 0.75, and low if it is higher than 0.75. Based on these values, it is observed that soil organic carbon across the Chad area has a medium spatial dependency level, with the nugget-to-sill ratio value of 0.37 falling between 0.25 and 0.75. The SOC spatial distribution map generated on the basis of SoilGrids data is shown in Fig. 4, and the prediction standard error map in Fig.5. The analysis of both figures shows that the southern portion of Chad is richer than the northern portion of the country. This could be justified by the fact that the northern portion is desert, while the southern portion is greener and used for agricultural activities.



3.2.1. Spatial Distribution of Predicted SOC Map

The spatial distribution map of SOC generated from the Gaussian variogram model shows higher SOC concentrations in the southern portion of Chad, reflecting the region's higher soil fertility level and agricultural activity. This distribution pattern is consistent with the area's known climatic and soil conditions, where more humid conditions favor higher organic carbon content. The accuracy of the map was further confirmed through cross-validation, where the predicted SOC values closely matched the observed data at known sample locations.

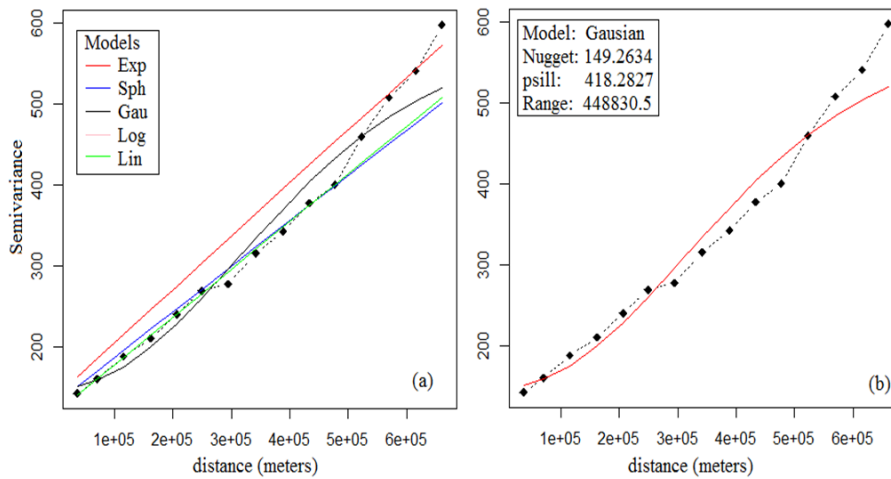


Fig 3: (a) Semi-variogram models applied on the SOC variogram cloud; (b) Best-fit semi-variogram model for the spatial distribution of soil organic carbon in Chad.

Table2: Choice of variogram model

Models	RMSE	ME	MAE
Exponential	36.49	29.40	32.80
Spherical	35.25	-11.80	22.32
Gaussian	27.84	-3.35	20.95
Logarithmic	6428.91	6428.52	6428.51
Linear	32.22	-14.11	18.24

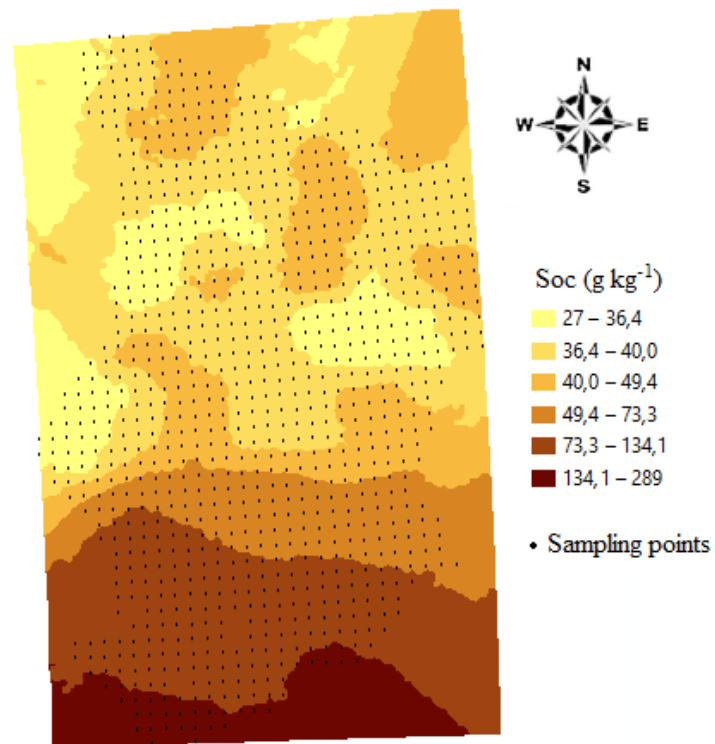


Fig. 4: Predicted spatial distribution map of soil organic carbon (SOC) in Chad, based on SoilGrids data and a fitted variogram.

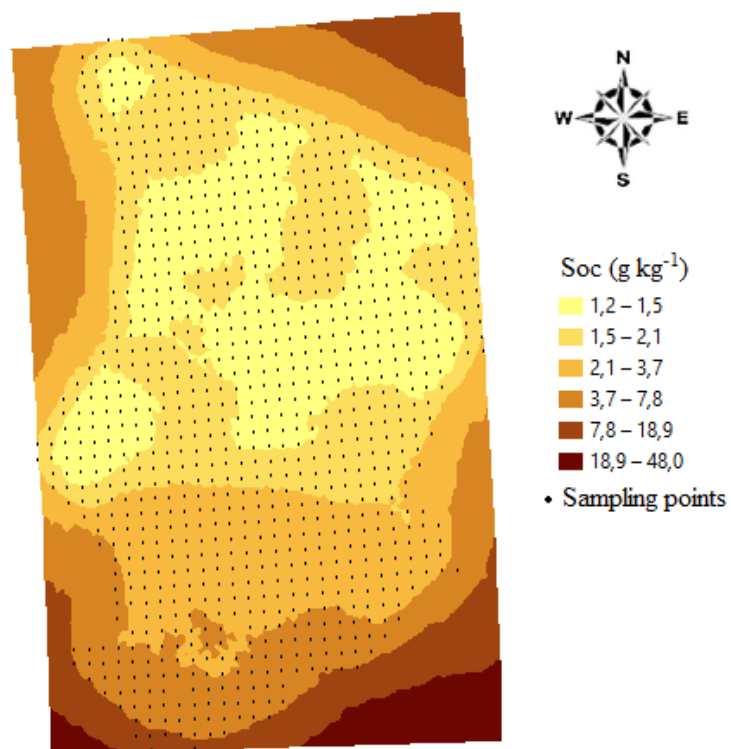


Fig. 5: Standard error map of SOC spatial prediction in Chad.

#### **4. CONCLUSION**

This study has successfully illustrated the medium spatial autocorrelation level of soil organic carbon (SOC) across different regions of Chad, as characterized by a nugget-to-silt ratio of 0.37. This ratio indicates a significant but not overwhelming influence of spatial factors on SOC distribution, suggesting that both local site conditions and broader regional factors play roles in shaping SOC levels. Identification of such a level of spatial autocorrelation is critical for improving the precision of SOC predictions and guiding soil management practices that can be specifically tailored to different areas. The SOC spatial distribution maps generated by this study serve as valuable tools for stakeholders in agricultural planning and environmental management. By providing a clearer picture of SOC variability, these maps allow for more targeted soil-sampling efforts, reducing the need for extensive field surveys and thus optimizing both the cost and efficiency of soil management practices. Furthermore, the maps can be used to identify potential areas for soil degradation or improvement, thus guiding interventions that can improve soil health and agricultural productivity. Additionally, the methodology applied in this study offers a replicable model for updating SOC maps within shorter timespans, an essential feature in the context of rapid environmental change and agricultural intensification. The ability to quickly update these maps ensures that they remain relevant and useful for ongoing soil management and conservation efforts. Our findings from this study not only contribute to the scientific understanding of SOC dynamics in Chad, but also provide practical tools and approaches that can significantly impact sustainable agricultural practices and environmental conservation strategies. Moving forward, it is recommended that further research be conducted to refine SOC prediction models and to explore the impacts of different land use practices on SOC levels, ultimately supporting the development of more resilient agricultural systems and improved land management policies.

#### **5. REFERENCES**

- Addise, T., Bedadi, B., Regassa, A., Wogi, L., & Feyissa, S. (2022). Spatial variability of soil organic carbon stock in Gurje Subwatershed, Hadiya Zone, Southern Ethiopia. *Applied and Environmental Soil Science*, 2:1-12. DOI:10.1155/2022/5274482
- Akpan, S. I., Odeh, I. O., Bishop, T. F., Hartemink, A. E., & Amapu, I. Y. (2016). Total soil organic carbon and carbon sequestration potential in Nigeria. *Geoderma*, 271, 202-215.
- Ba, A. *Chad Agriculture Production Support Project: P126576 - Implementation Status Results Report: Sequence 10*. The World Bank; 25 April 2017.
- Bhunja, G. S., Shit, P. K., & Maiti, R. (2018). Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). *Journal of the Saudi Society of Agricultural Sciences*, 17(2), 114-126.
- Boubehziz, S., Khanchoul, K., Benslama, M., Benslama, A., Marchetti, A., Francaviglia, R., & Piccini, C. (2020). Predictive mapping of soil organic carbon in Northeast Algeria. *Catena*, 104539.

- Cerri, C. C., Bernoux, M., Arrouays, D., Feigl, B. J., & Piccolo, M. C. (2000). Carbon stocks in soils of the Brazilian Amazon. In Lal R., Kimble J. M., Stewart B. A. (Eds.), *Global Climate Change and Tropical Ecosystems* (pp. 33-50). CRC/Lewis Publishers, Boca Raton.
- Chabala, L. M., Mulolwa, A., & Lungu, O. (2017). Application of ordinary kriging in mapping soil organic carbon in Zambia. *Pedosphere*, 27(2), 338–343.
- Du, M. Y., Fan, S. H., Liu, G. L., Feng, H. Y., Guo, B. H., & Tang, X. L. (2016). Stoichiometric characteristics of carbon, nitrogen, and phosphorus in *Phyllostachys pubescens* forests of China. *Chinese Journal of Plant Ecology*, 40, 760–774.
- Fromm, S. F. von, Hoyt, A. M., Lange, M., Acquah, G. E., Aynekulu, E., Berhe, A. A., & Shepherd, K. D. (2021). Continental-scale controls on soil organic carbon across sub-Saharan Africa. *Soil*, 7(2), 305-332.
- Guèye, B. (2003). L'agriculture familiale en Afrique de l'Ouest, concepts et enjeux actuels. *Revue Afrique Contemporaine*.
- Jarnevich, C. S., Hayes, M. A., Fitzgerald, L. A., Yackel Adams, A. A., Falk, B. G., Collier, M. A., & Reed, R. N. (2018). The distribution of tegu lizards in native South America and potentially invasive ranges is modeled. *Scientific Reports*, 8(1), 1-12.
- Jennings, E., Allott, N., Pierson, D. C., Schneiderman, E. M., Lenihan, D., Samuelsson, P., & Taylor, D. (2009). Impact of climate change on phosphorus loading from a grassland catchment: Implication for future management. *Water Resources*, 43, 4316–4326.
- Li, Y., Niu, S. L., & Yu, G. R. (2016). Aggravated phosphorus limitation on biomass production under increasing nitrogen loading: a meta-analysis. *Global Change Biology*, 22, 934–943.
- Mahmoudzadeh, H., Matinfar, H. R., Taghizadeh-Mehrjardi, R., & Kerry, R. (2020). Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*, 21, e00260.
- Marklein, A. R., & Houlton, B. Z. (2012). Nitrogen inputs accelerate phosphorus cycling rates across a wide variety of terrestrial ecosystems. *New Phytologist*, 193, 696–704.
- Nagle, N. N. (2010). Geostatistical smoothing of areal data: Mapping employment density with factorial kriging. *Geographical Analysis*, 42(1), 99-117.
- N'Guessan, A. A. B., & Bernard, D. K. (2012). Changements climatiques, agriculture et sécurité alimentaire en Afrique subsaharienne : le cas de la Côte d'Ivoire.
- Owusu, S., Yigini, Y., Olmedo, G. F., & Omuto, C. T. (2020). Spatial prediction of soil organic carbon stocks in Ghana using legacy data. *Geoderma*, 360, 114008.
- Razakarison, N. V. R. (2013). Variabilité climatique et insécurité alimentaire (Doctoral dissertation, Université d'Antananarivo).
- Ryu, J. S., Kim, M. S., Cha, K. J., Lee, T. H., & Choi, D. H. (2002). Kriging interpolation methods in geostatistics and DACE model. *KSME International Journal*, 16(5), 619-632.
- Sreenivas, K., Dadhwal, V. K., Kumar, S., Harsha, G. S., Mitran, T., Sujatha, G., ... & Ravisankar, T. (2016). Digital mapping of soil organic and inorganic carbon status in India. *Geoderma*, 269, 160-173.
- Tripler, C. E., Kaushal, S. S., Likens, G. E., & Walter, M. T. (2006). Patterns in potassium dynamics in forest ecosystems. *Ecological Letters*, 9, 451–466.
- Wang, Y. Q., & Shao, M. A. (2013). Spatial variability of soil physical properties in a region of the Loess Plateau of PR China subject to wind and water erosion. *Land Degradation and Development*, 24(3), 296–304.