SPATIAL PREDICTION OF SOIL ORGANIC CARBON IN PANAMARAM BLOCK, KERALA - A COMPARITIVE STUDY OF VARIOUS TECHNIQUES

S. R. Bhavana[1], T. Radhakrishnan[2], Anju S[3]

[1]Department of Geography, Nirmala College for Women, Coimbatore, Tamil Nadu, India

[2] School of Digital Sciences, Digital University Kerala, Thiruvananthapuram, Kerala, India

[3]Centre for Geospatial Analytics, School of Digital Sciences, Digital University Kerala, Thiruvananthapuram, Kerala, India

* Corresponding author

doi: https://doi.org/10.21467/proceedings.179.18

ABSTRACT

Soil Organic Carbon (SOC) is the carbon that is contained within soil organic matter and it plays a main role to soil health, fertility and ecosystem services, which includes food production – making its preservation and restoration essential for sustainable agriculture development. SOC plays an important role in the global carbon cycle and, consequently, in climate change mitigation and adaptation. Sustainable management of SOC in the soil demands the scientific knowledge of its spatial distribution. In this study, we used five spatial interpolation algorithms to spatially predict the SOC in the Panamaram Block of Wayanad District in Kerala. Five spatial interpolation algorithms, including two variants of Kriging, were applied in the soil nutrient data. The performance of the prediction algorithms was compared using prediction accuracy parameters such as R^2, RMSE and MAE. Experimental results revealed that predictions generated by Inverse Distance Weighting (IDW) exhibit the highest accuracies with an R^2 value of 99.25%, followed by the Radial Basis Polynomial Interpolation Method (67.03%), Ordinary Kriging (26.79%), Local Polynomial Interpolation (18.96%), and Simple Kriging Method (13.07%), respectively. This research contributes valuable insights into understanding the spatial distribution of Soil Organic Carbon which is pivotal for informed land management and environmental conservation strategies.

Keywords: Soil Organic Carbon, Panamaram Block

1. INTRODUCTION

The spatial variability of Soil Organic Carbon (SOC) is an essential indicator of soil quality and a significant component of carbon reserves in land-based ecosystems. It plays a crucial role in ecosystem modelling, environmental prediction, precision farming, and the management of natural resources. Understanding the spatial distribution of SOC is crucial for evaluating soil fertility and crafting effective agricultural environmental management strategies. Proper management of SOC is key to sustainable agricultural development, requiring detailed information about its spatial and temporal behaviour across different regions [1]. However, assessing SOC is intrinsically expensive and time-intensive, especially during the initial soil sampling phase. SOC, a vital indicator of soil productivity, encompasses the organic fraction of undecayed plant and animal residues within the soil. It bolsters the sustainability of cropping systems by improving soil physical properties (such as texture, structure, density, and water-holding capacity), chemical properties (like nutrient availability and mitigation of aluminum

Proceedings DOI: 10.21467/proceedings.179; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-984081-7-4

toxicity), and biological properties (including nitrogen mineralization, nitrogen fixation, and microbial biomass). [2]

Spatial interpolation techniques are valuable tools for predicting the values of primary variables at unsampled points within the same region (interpolation) and beyond the region covered by existing observations (extrapolation). These methods are essential for estimating spatially continuous data and have been applied across various Earth surface disciplines [3]. With the increasing application of spatial interpolation methods, more than ten models have been developed across different fields, raising concerns about their accuracy. Extensive literature has been published comparing the performance of these interpolators using specific datasets. The choice of interpolation technique depends on the available data, spatial and temporal scales, desired outcomes, and required accuracy. Commonly used interpolation techniques are Inverse Distance Weighting (IDW), Kriging (Ordinary and Simple), Radial Basis Functions, Local Polynomial Interpolation [4]. The objective of this study is to spatially predict the variation of Soil Organic Carbon from point samples and compare the performance of various prediction techniques.

2. STUDY AREA AND SAMPLE DATA

The Panamaram Block of Wayanad district in Kerala State (see Figure 1) has been chosen for this study. Wayanad is a district, which is situated in the North-East part of Kerala with its administrative headquarters in the Municipality of Kalpetta. Wayanad was established on November 1, 1980, and it is the 12th district of Kerala, formed by carving out areas from Kozhikode and Kannur districts. Geographically, the district lies between 11°27' and 15°58' north latitude and 75°47' and 70°27' east longitude, encompassing an area of 2,132 sq km, of which 885.92 sq km is forested. The name Wayanad was originated from "Vayal Nadu," meaning the land of paddy fields. It is a scenic plateau located at an altitude ranging between 700 and 2100 meters above the mean sea level. It is nestled among the mountains of the Western Ghats, on the eastern side of North Kerala, bordering the states of Tamil Nadu and Karnataka. Wayanad is endowed with abundant water resources, making it an ideal region for paddy cultivation, from which the district derives its name. The climate varies with temperatures rising up to 35°C during the hot weather and dropping to 10°C during the cold weather, accompanied by an average annual rainfall of 300 mm. The cool climate favours the cultivation of plantation crops such as coffee, tea, and spices, which are extensively grown across the district. The high altitude of Wayanad provides suitable soil conditions for the cultivation of perennial crops and spices. The district is renowned for its coffee, tea, pepper, cardamom, and rubber plantations, which thrive on the hills of the area. The climatic conditions of Wayanad contribute to its agricultural productivity, offering good yields and high resistance to pest and disease attacks. Integrated farming with conventional crops and exotic fruit plants is particularly suitable for Wayanad farmers, especially in the context of climate change. This integration ensures a sustainable income for farmers; if one crop fails, others can compensate, providing economic stability. [5]. Given the diverse weather patterns, climatic conditions, vegetation, terrain, soil conditions, and the abundant natural resources relatively untouched by human intervention, this study aims to improve agricultural prospects in the Panamaram block. The central Block Panchayat of Panamaram was selected to investigate the Soil Organic Carbon (SOC) content in the soil nutrients of the area. Soil samples from the Panamaram block were collected through field surveys, and their SOC values were determined in the lab. Approximately 2900 sample points, scattered randomly across the block, were analyzed for this purpose (as shown in Figure 1). The minimum and maximum value of SOC range from 0 to 7.03ppm.

Proceedings DOI: 10.21467/proceedings.179
ISBN: 978-81-984081-7-4

Series: AIJR Proceedings
ISSN: 2582-3922

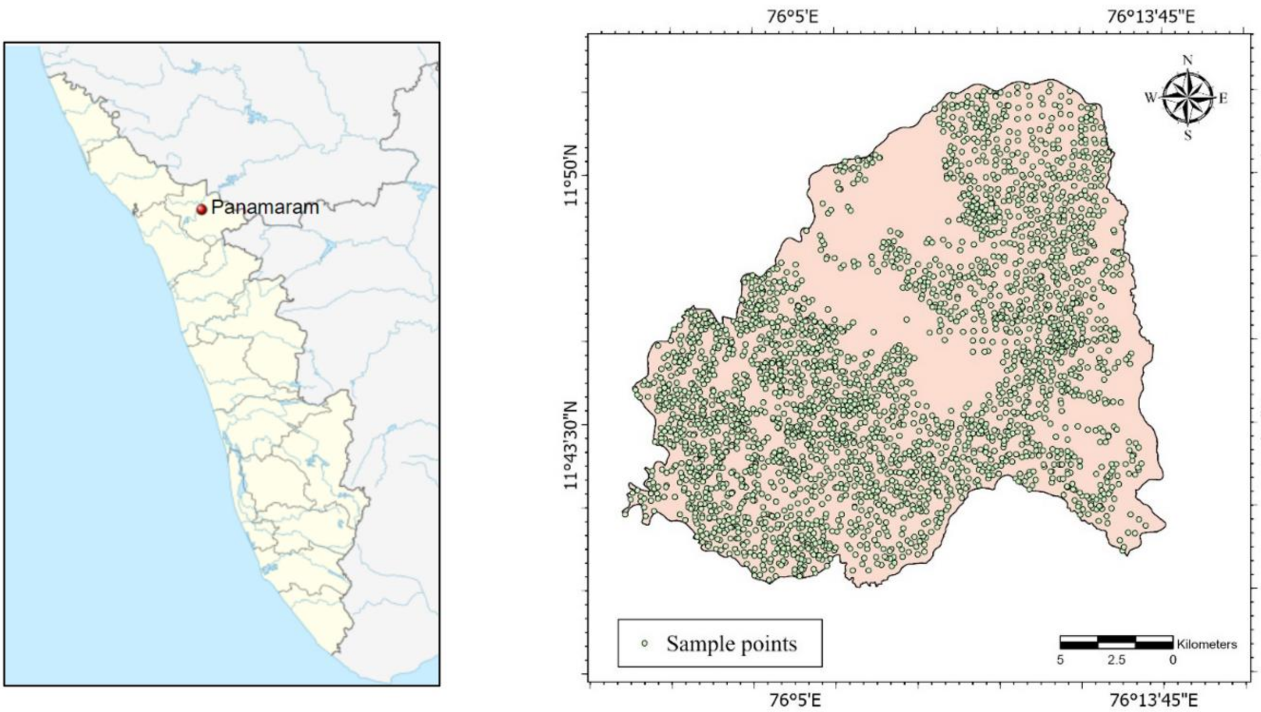

Figure 1 : Study Area Boundary

3. METHODOLOGY

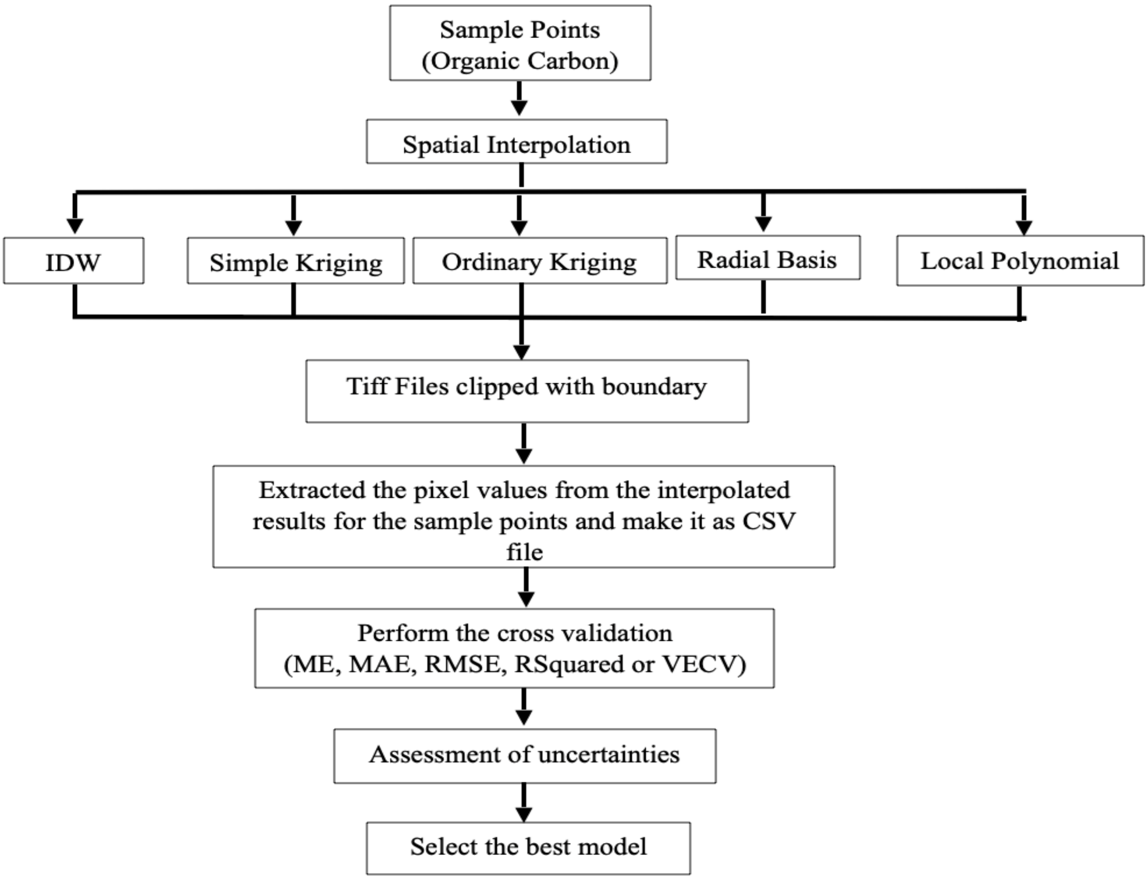

Figure 2 : Workflow of the study

Proceedings DOI: 10.21467/proceedings.179
ISBN: 978-81-984081-7-4

Series: AIJR Proceedings
ISSN: 2582-3922

3.1 Data Collection and Preparation

Soil samples from the Panamaram block were collected and recorded in an Excel sheet, including the coordinates (longitude and latitude) and the SOC values for each point. The data were converted into point locations and visualized to identify any incorrect coordinates or misplaced points. The Coordinate Reference System of the sample data is projected coordinate system, EPSG 32643. Exploratory spatial data analysis was also conducted. Mean, Median and Variance of the SOC values were calculated. The distribution of the SOC values is visualized by plotting the histogram and boxplots and also identified the outliers to ensure the normal distribution of the data points

For accurate prediction results, the sample data points should ideally follow a normal distribution. Any identified outliers were removed to achieve a normal distribution of the data points. In the present study, both deterministic and geostatistical interpolation techniques were employed to predict Soil Organic Carbon (SOC) levels. The deterministic techniques create surfaces from the sample points, while geostatistical techniques apply the statistical properties of the measured points. After preprocessing of data, the different interpolation techniques like IDW, Simple Kriging, Ordinary Kriging, Radial Basis Function, and Local Polynomial Interpolation were applied. IDW is a deterministic method that weighs the sample points by their distance to the cell being estimated with closer points having more influence [4]. Both Ordinary Kriging and Simple kriging uses distance and spatial arrangement of points to estimate values, suitable for the data with spatial correlation [1]. The Radial Basis Functions generates surfaces that pass through each measured point [6] The local polynomial Interpolation provides a local less smooth surface [7]. The different tools available in ArcGIS Pro software is used for performing the interpolation of the sample points (as can be seen in Figure 2)

4. RESULT

The different interpolation techniques that were used is compared through an accuracy assessment in which the difference between the known data and the predicted data is examined using the mean absolute error (MAE), root mean squared error, Mean error and R squared error or Vec value. Mean Absolute Error (MAE) is calculated by averaging the absolute differences between the original and predicted values over a dataset, making it the simplest regression error metric to comprehend. An MAE of 0 signifies perfect model predictions. Mean Error (ME), on the other hand, is derived by averaging the squared differences between the original and predicted values [8]. Root Mean Squared Error (RMSE) converts the error metric back to the same units as the original output, aiding in easier interpretation [9]. Since both MAE and MSE can range from 0 to positive infinity, interpreting model performance becomes more challenging as these values increase. R squared, or Vecv, is an accuracy measure that is independent of unit/scale, data mean, and variance. It unifies error measures through various methods and is interpreted as a percentage ranging from 0 to 1. Higher values indicate a better model.

Proceedings DOI: 10.21467/proceedings.179
ISBN: 978-81-984081-7-4

Series: AIJR Proceedings
ISSN: 2582-3922

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \quad (1)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \quad (2)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \quad (4)$$

Where,

$\hat{y}$ – predicted value of y

$\bar{y}$ – mean value of y

For performing the accuracy assessment, the prediction results are converted to a comma separated value file (the test results are summarized in Table 1) which contains the predicted value for SOC of all the interpolation methods. This csv file is then opened in R studio and performed the accuracy assessment. The result obtained for each method is again converted to a csv file which contain ME value, MAE value, RMSE value and R Squared Value. The table below shows the accuracy result of each interpolation technique.

Table 1: Accuracy assessment of different spatial interpolation method

S.No.	Method	ME	MAE	RMSE	VECV
1.	IDW	-0.000	0.0195	0.0673	99.254
2.	Ordinary Kriging	0.001	0.5001	0.6671	26.799
3.	Simple Kriging	0.117	0.5305	0.7269	13.071
4.	Radial Basis	0.002	0.3329	0.4477	67.034
5.	Local Polynomial	0.002	0.5258	0.7019	18.958

From the above accuracy table (1), it is inferred that even though none of the methods are giving out same values. The ideal values for correct VECV and ME should be closer to 1 and 0 respectively. The higher the VECV digits (closest to 1) and the lower the ME digits (closest to 0), the most accurate method would be to get a solution. Thus, according to the data in the table above, IDW method is the most accurate as VECV is highest (99.254 – i.e. closest to 1) here along with ME having the lowest value (-0.000 – i.e. closest to 0). It is followed by Radial Basis, Ordinary Kriging, Local Polynomial Interpolation and Simple Kriging, in that order. We

Proceedings DOI: 10.21467/proceedings.179
ISBN: 978-81-984081-7-4

Series: AIJR Proceedings
ISSN: 2582-3922

can infer that Simple Kriging has the least value in both the VECV as well as ME values and hence this form of kriging is the least accurate one for our data. This all is done as part of accuracy assessment.

4.1 Visualization

The prediction result generated after performing the spatial interpolation is in the form of raster surface. It is then saved to tiff file. These files were then extracted for the study area boundary to focus on the Panamaram block. Then tiff files are then analyzed for identifying the spatial pattern of SOC values and the high or low concentration areas. The prediction maps were generated for showing the spatial distribution of the SOC values across the study area. The predicted low SOC values which are less than 1 are represented by dark green and the high SOC values of greater than 2 by red color in the map (see Figure 3 - Figure 7 in prediction maps)

4.2 Prediction Maps.

The final prediction maps of SOC using different interpolation methods are shown as under :-

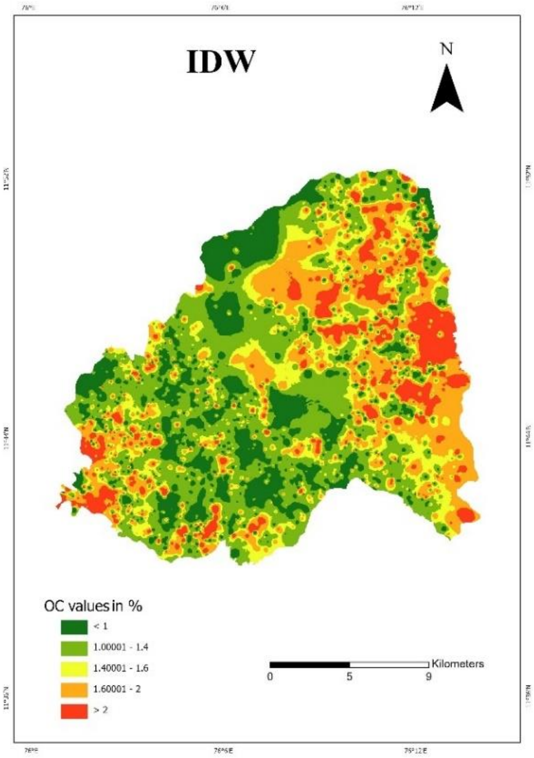

Figure 3 : Spatial Variability map for IDW method

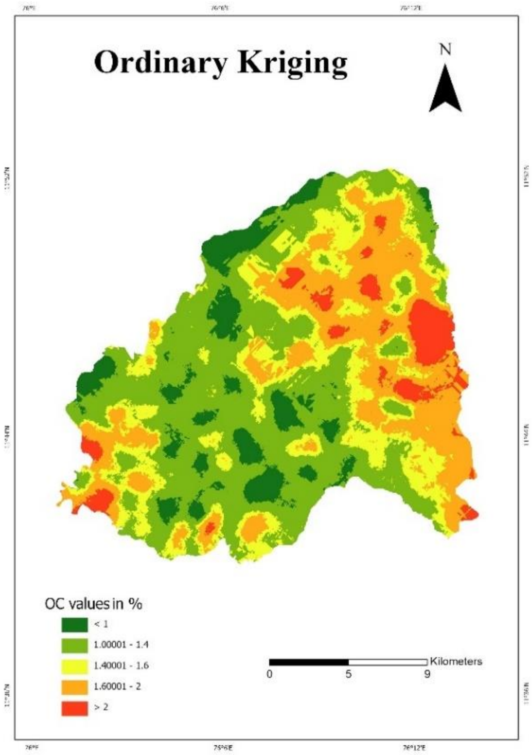

Figure 4 : Spatial Variability map Ordinary Kriging method

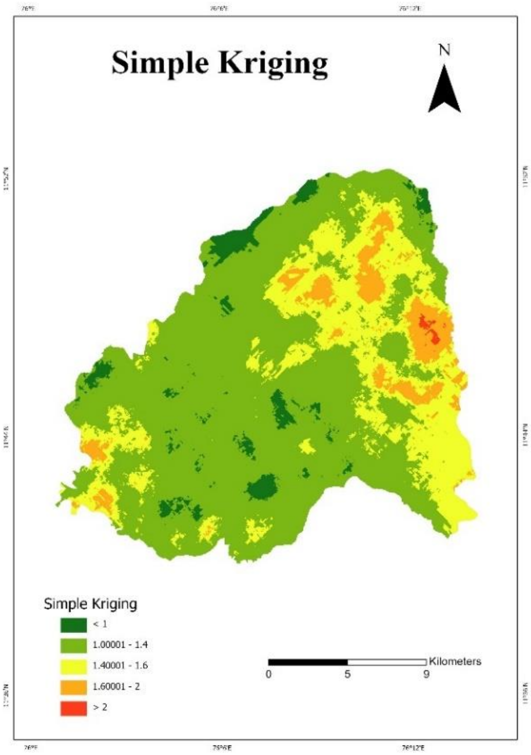

Figure 5 : Spatial Variability map for Simple Kriging method

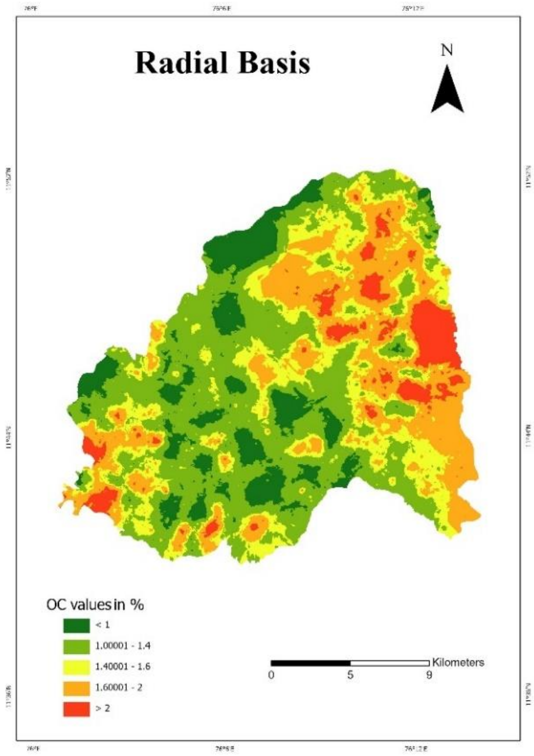

Figure 6 : Spatial Variability map for Radial Basis method

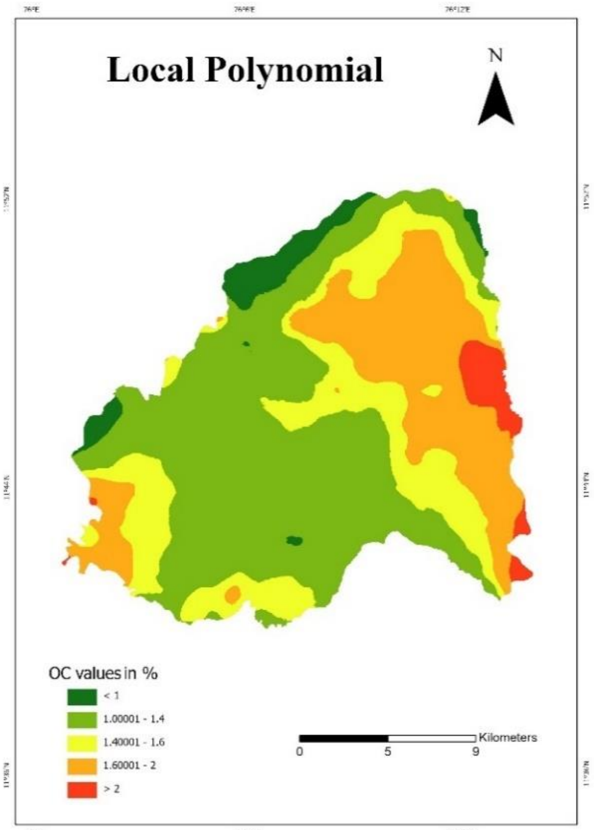

Figure 7 : Spatial Variability map for Local Polynomial method

Proceedings DOI: 10.21467/proceedings.179
ISBN: 978-81-984081-7-4

Series: AIJR Proceedings
ISSN: 2582-3922

After the comparison, the IDW Method map is the most accurate one (see Figure 3 for instance). Now when IDW is taken as the accuracy result interpreter, the Soil Organic Carbon (SOC) of Panamaram block can be found more where some sort of vegetation, that of a forest land or manmade cultivated area is present and found less in barren lands where no sort of vegetation is found as vegetated land would be more fertile and presence of the SOC can be seen. This data is sourced from the study done along with the base map interpretation of Panamaram block

5. CONCLUSION

In the above study, different spatial interpolation methods like Inverse Distance Weighting, Kriging, Local Polynomial and Radial basis function were employed to examine the spatial distribution of Soil Organic Carbon content in the Panamaram Block of Wayanad District, Kerala. Accuracy assessment was carried out using various regression models such as ME, MAE, RMSE, VECV. Although all the methods and various regression models used showed satisfactory results, the IDW method has been found most accurate for the Soil Organic Content Prediction Map. It is desirable to compare these results with other interpolation methods at different scale.

6. CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest related to the content of this manuscript. No financial or personal relationships have influenced the research or the interpretation of results. The authors have received no funding from any organizations or entities that could benefit from the publication of this manuscript, and there are no competing interests in terms of intellectual property, financial investments, or affiliations that might affect the objectivity of the research presented.

REFERENCES

[1] Gouri Sankar Bhunia, Pravat Kumar Shit, Ramkrishna Maiti (2016)- “Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC)”

[2] Nega Emiru and Heluf Gebrekidian (2006-2013) – “Effect of land use changes and soil depth on soil organic matter, total nitrogen and available phosphorous contents of soil in Senbat Watershed, Western Ethiopia.”

[3] Kingsley John, Isong abraham ,Ndiye Michael Kebonye, Prince Chapman Agyeman, Esther Okon Ayito and Ahado Samuel Kudjo (2021) – “Soil organic carbon prediction with terrain derivatives using geostatistics and sequential Gaussian simulation.”

[4] Arun Kumar Gumudaveli (2022) – “Accuracy assessment of various interpolation techniques used for generation of precise geoidal undulations in LiDAR projects.”

[5] en.wikipedia.org/wiki/Panamaram (2022).

[6] Fereydoon Sarmadian and Ghavammodin Zehedi Amiri (2014) – “Mapping of spatial variability of soil organic carbon based on radial basis functions method.”

[7] Bhunia G.S., Shit P.K and Maiti R. (2018) – “Comparison of GIS based interpolation methods for spatial distribution of soil organic carbon - Journal of Saudi Society of Agricultural Sciences.”

[8] Yashmeet Singh (2019) – “Data tech notes: Machine learning regression model accuracy.”

[9] Shweta Acharya (2021) – “A simple guide to evaluation metrics.”