# Prediction of Soil colour using vis-NIR spectroscopy and machine learning models

## Abstract:

Soil colour is a critical indicator of soil properties and conditions, influencing various agronomic and environmental factors. A total of 2216 surface soil samples (0-15 cm) were collected from the Kymore Plateau and Satpura Hill zone of Madhya Pradesh, using Global Positioning System (GPS) for precise location. Soil colour parameters were measured in the field using the Munsell soil colour chart, while chemical analysis was conducted in the laboratory following standard procedures. Additionally, spectra of the soil samples were recorded using a spectroradiometer under laboratory conditions.

The results showed that the soil colour hues ranged from 10R, 10YR, 2.5Y, 2.5YR, 5Y, 5R, 5YR, to 7.5YR, Values and Chroma varied from 2 to 7 and 1 to 8, respectively. Correlation analysis revealed negative correlations between the RGB components and organic carbon, with r values of -0.114**, -0.071**, and -0.101* for R, G, and B, respectively. Polynomial models showed the best fit for the relationship between the value and chroma of the colour parameters and soil organic carbon (SOC), with equations $Y = 0.086x^2 - 0.860x + 7.528$ ($R^2 = 0.982$) and $Y = 0.018x^2 - 0.249x + 6.126$ ($R^2 = 0.948$), respectively. A linear relationship was observed between chroma and available phosphorus (P), with the equation $Y = -0.873 + 13.92$ ($R^2 = 0.922$).

In addition, machine learning models, including PLSR, SVM, Random Forest, ANN, XGBoost, LightGBM, CatBoost, and ELM algorithms, were used to predict soil colour parameters. Among these, the Random Forest and XGBoost models demonstrated the best performance in predicting soil colour parameters (L*, a*, b*, R, G, and B), with model accuracies of 83.6%, 80.9%, 83.0%, 84.3%, 83.7%, and 83.4%, respectively. soil colour variation depicted in the maps generated using GIS can also serve as covariates for mapping, offering comprehensive insights into the soil's properties.

**Key words:** Soil colour, Spectroscopic, Machine learning models, Munsell colour chart, GPS

### 1.0 Introduction

The soil colour play crucial role in soil science, frequently documented in soil profile descriptions due to its role as an initial indicator of various soil conditions and characteristics. Observations made in field studies reveal that soil colour is not only readily discernible but also closely associated with numerous other soil properties (Schmidt et al., 2021). It functions effectively as a rapid assessment tool for parameters including drainage class, soil classification, and organic carbon content, all of which may be influenced by the application of variable rate fertilizers and pesticides, as indicated by Moritsuka et al. (2019). Established correlations have been noted between soil colour and soil texture, water content, iron content, and organic carbon levels. Generally, darker surface soils correlate with elevated organic matter content, which implies fertile conditions that are conducive to plant growth. Such soils are often perceived to possess favorable qualities, including good drainage, adequate aeration, high nitrogen content, and lower susceptibility to erosion. Conversely, lighter-coloured soils are typically associated with contrasting characteristics.

The characterization of soil colour is commonly performed utilizing the Munsell colour system, a method that has gained widespread acceptance among soil surveyors for classification purposes, thereby serving as a fundamental tool in the field of soil descriptions. Although the Munsell system remains dominant, alternative colour models, such as the RGB (Red, Green, Blue) system, are frequently employed in digital contexts. The process of characterizing soil colour through Cartesian systems can be conducted directly or achieved through the conversion of Munsell data into these alternative systems via lookup tables or statistical regression, as outlined by Dominguez Soto et al. (2012). Moreover, advancements in technology have significantly improved the efficient and precise acquisition of digital soil colour data, facilitating analyses that are more consistent and accurate in comparison to traditional visual methods.

Despite multiple attempts to compile regional maps of soil colour, a comprehensive map of topsoil colour has yet to materialize, as noted by Poppie et al. (2020). Mapping initiatives have encountered limitations, often relying on interpolations of point-based soil observations from regions in Australia (Viscarra-Rossel et al., 2010), China (Liu et al., 2020), and other locations (Soils, 2023). Due to the discrete nature of soil colour, it does not inherently reflect the continuous spatial variability that exists within soils. As a result, interpolation maps generated from point-based data heavily depend on field observations collected at specific times, possibly failing to capture the broader and dynamic aspects of soil colour changes, as emphasized by Liu et al. (2020). The absence of fine-scale mapping of surface soil colour within the study area underscores the necessity for such efforts to effectively monitor ongoing soil conditions.

Soil colour is a fundamental property that provides insights into soil composition, organic matter content, and overall health. The darker soils have higher SOC than lighter soils. This is important for assessing carbon sequestration, managing and soil health (Liu et al. (2020). Traditional soil analysis often requires sample collection, preparation, and laboratory testing, which can be time-consuming and expensive. Vis-NIR spectroscopy offers a non-destructive, rapid, and cost-effective alternative, enabling real-time or field-based soil analysis without the need for extensive sample handling (Sawut et al. 2018 and Lin et al. 2019). Machine learning models can process large amounts of spectral data quickly, providing timely predictions of soil colour and properties.

## 2.0 Materials and methods

**2.1 Study area:** The Kymore plateau and Satpura hills zone selected for study which comprise of Rewa, Satna, Panna, Jabalpur, Seoni, Katni, Sidhi and Singrauli districts covering 49.97 lakh ha (16.26%) area of the state. The region has a relatively high proportion of waste and uncultivated lands-about 21%. Another about 22% of the land is under forest cover. Only 37% is cultivated. Irrigation facilities are very poor as only about ten per cent of the cultivated land is irrigated. Mixed red and medium black soils are mainly in the region. In this region, annual rainfall is in the range of 1100 to 1400 mm and wheat and rice are main crops. GPS based 2216 surface soil samples (0-15 cm) were collected from from farmer's fields during the off season of 2022-2023.

### 2.2 Sample collection and processing

At each sampling point, a 1-kilogram composite soil sample, which accurately represents the area, was collected and recorded in a correctly labeled sample bag. The collected samples were air dried in the shade, crushed using a wooden log to break up clumps and aggregates, and any visible root fragments were removed. Each sample was then passed through a 2 mm sieve and samples were air-dried, ach sample was divided into two portions, one for laboratory chemical analysis and the other one for spectral measurements. The soil pH EC SOC P,K and S were analsysed using standard procedures.

### 2.3 Soil colour

The colour parameters of soil (i.e., Hue, Value and Chroma) were recorded by Munsell colour chart (1994 Revised Edition). By placing the sample directly behind the colour sheets separating the nearest matching colour chips, the colour of the soil samples was matched. Hue, Value, and Chroma were recorded and also converted from HVC to L* b* a* R G B colour parameter as described by Kirillova *et al*., (2018).

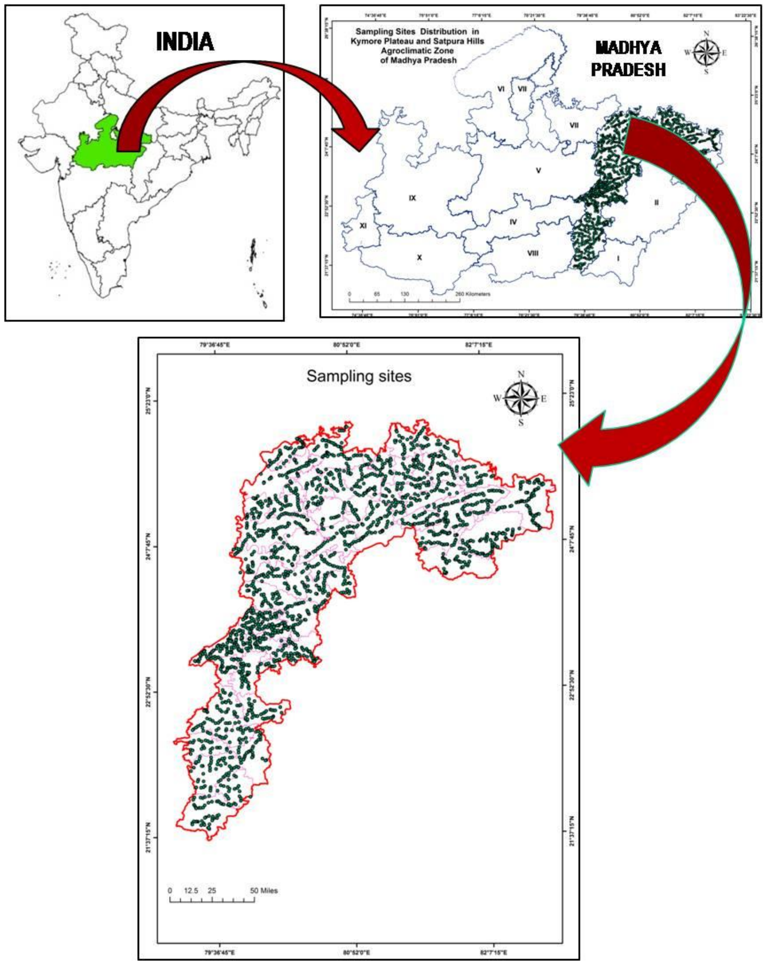



**Fig. 1 Location map of study area**

### 2.4 Spectral reflectance measurements

The spectra reflectance of soil samples recorded with the help of Spectroradiometer (RS-3500). Air-dried samples were placed evenly on a rectangular black disk of 5 cm diameter and 2 cm depth by tapping the tray on a table to ensure a smooth surface. The soil-filled rectangular black disk was kept on the dark background of a table and light reflectance was measured. Reflectance measurements were taken under dark room conditions and colour estimates were made in the 350 to 2500 nm wavelength region of the spectrum.

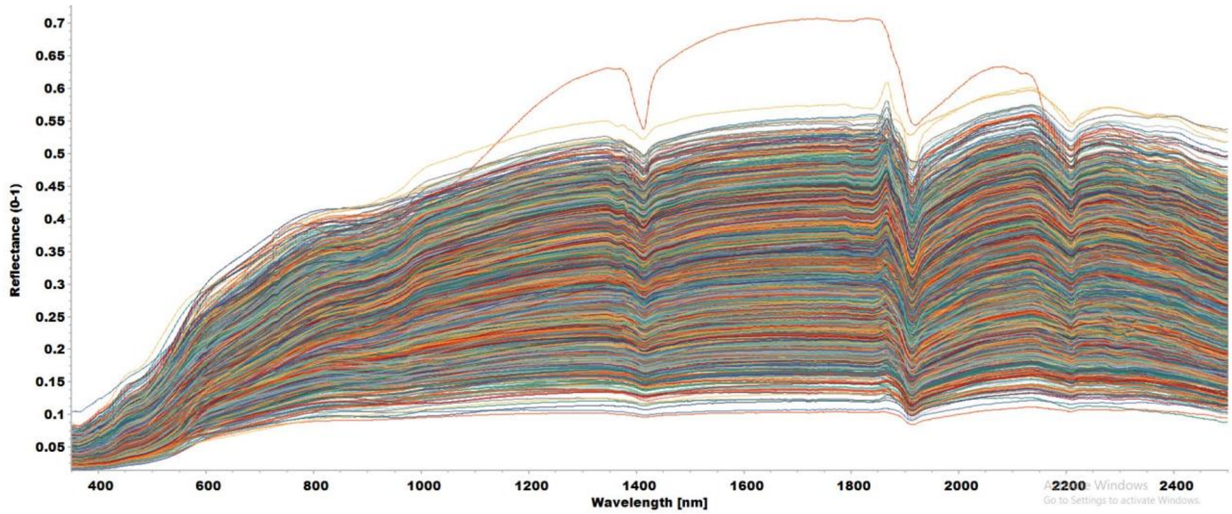



Fig 2. Spectral signature of different soils

**2.5 Model building**

The hyperspectral data were analyzed to determine RGB and correlated with Munsell Soil colour measurement using correlation techniques. The PLSR, SVR, ANN, Random Forest, XGBoost, Light GBM, Cat Boost machine learning algorithms to develop the model for the prediction of soil colour parameters, the whole data set (n=2217) was divided 1:5 ratios (every fifth sample used for testing dataset) into two datasets *viz.*, training dataset (n=1774) and testing dataset (n=443) in the 80:20 ratios. The training dataset was used to derive the spectral model and the testing dataset was applied to check the predictive performance of the newly developed model. The model's performance was evaluated using testing dataset metrics such as $R^2$, RMSE and RPD values.Mapping of soil colour parameter mapped using GIS

**3.0 Result and discussion**

**3.1 Variation of soil colour**

The Munsell soil colour parameters specifically Hue, Value, and Chroma, in the study area, as represented in Table 1 and depicted in Figures 3, 4, and 5, revealed significant variation across the samples collected from different districts, such as Jabalpur, Ktani, Panna, Rewa, Satna, Seoni, Sidhi, and Singrauli. The hues recorded were predominantly from the ranges of 10R, 10YR, 2.5Y, 2.5YR, 5Y, 5R, 5YR, and 7.5YR. The value ranged from 2 to 7, and chroma ranged from 1 to 8. These observations are consistent with findings of Foth, 1990; Brady and Weil, 2008 who have noted considerable soil colour diversity in tropical and subtropical regions, often influenced by factors such as mineral content, moisture, and organic matter.

The distribution of Hue in the samples was as follows: 10R (0, 96, 241, 12, 0, 19, 3, and 8), 10YR (0, 86, 75, 4, 0, 0, and 48), 2.5Y (0, 82, 132, 22, 0, 10, 9, and 46), 2.5YR (0, 46, 200, 12, 0, 0, 6, and 20), 5R (0, 71, 151, 12, 3, 4, 8, and 51), 5Y (4, 89, 131, 28, 0, 0, 10, and 38), 5YR (1, 61, 84, 16, 0, 0, 5, and 43), and 7.5YR (0, 80, 60, 7, 0, 0, 18, and 57). The number of samples observed within these ranges varied, with the majority of samples falling within the values of 2, 3, 4, 5, 6, and 7, indicating a broad range of soil brightness and tonal variations across the districts. This variability in soil colour, especially in relation to value and chroma, can be attributed to several factors, including differences in soil texture, mineral composition, and organic content Dinesh *et al.* (2017) and Arya *et al.* (2024). As noted by Schwertmann and Taylor (1989), the chroma and value of soil colour are often associated with the presence of iron oxides, which are influenced by the soil's drainage, aeration, and moisture content.

**Table: 1 Spatial distribution of soil colour in Kymore plateau and Satpura hills zone**

| Kymore plateau and Satpura hills | n | Hue | | | | | | | | Value | Chroma |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10R | 10YR | 2.5Y | 2.5YR | 5R | 5Y | 5YR | 7.5YR | | |
| **Jabalpur** | 379 | 0 | 96 | 241 | 12 | 0 | 19 | 3 | 8 | 2-7 | 1-8 |
| **Katni** | 220 | 0 | 86 | 75 | 4 | 0 | 0 | 7 | 48 | 2-7 | 2-6 |
| **Panna** | 301 | 0 | 82 | 132 | 22 | 0 | 10 | 9 | 46 | 2.5-7 | 1-8 |
| **Rewa** | 284 | 0 | 46 | 200 | 12 | 0 | 0 | 6 | 20 | 2.5-6 | 1-6 |
| **Satna** | 300 | 0 | 71 | 151 | 12 | 3 | 4 | 8 | 51 | 2-7 | 1-8 |
| **Seoni** | 300 | 4 | 89 | 131 | 28 | 0 | 0 | 10 | 38 | 2-6 | 2-6 |
| **Sidhi** | 210 | 1 | 61 | 84 | 16 | 0 | 0 | 5 | 43 | 2-6 | 1-8 |
| **Singrauli** | 222 | 0 | 80 | 60 | 7 | 0 | 0 | 18 | 57 | 2.5-6 | 1-8 |
| **Total** | **2216** | **5** | **611** | **1074** | **113** | **3** | **33** | **66** | **311** | **2-7** | **1-8** |

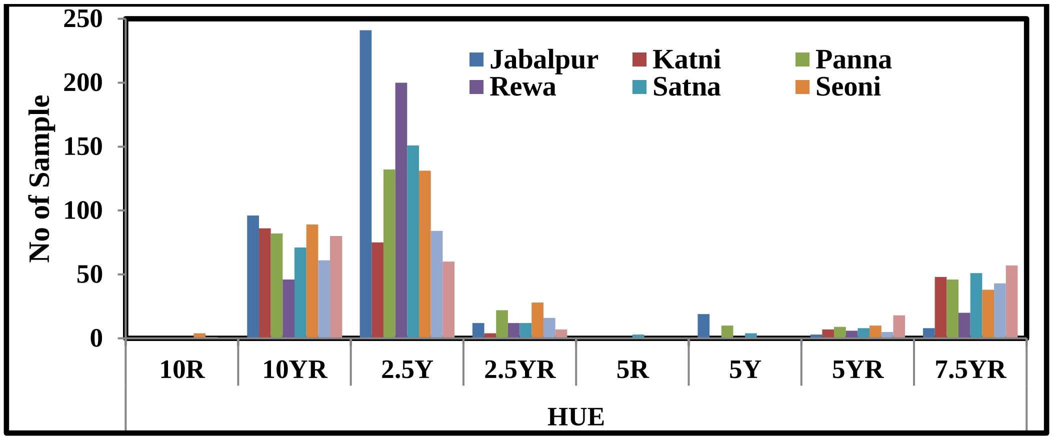



**Fig. 3 Frequency distribution of Hue in Kymore plateau and Satpura hills zone**

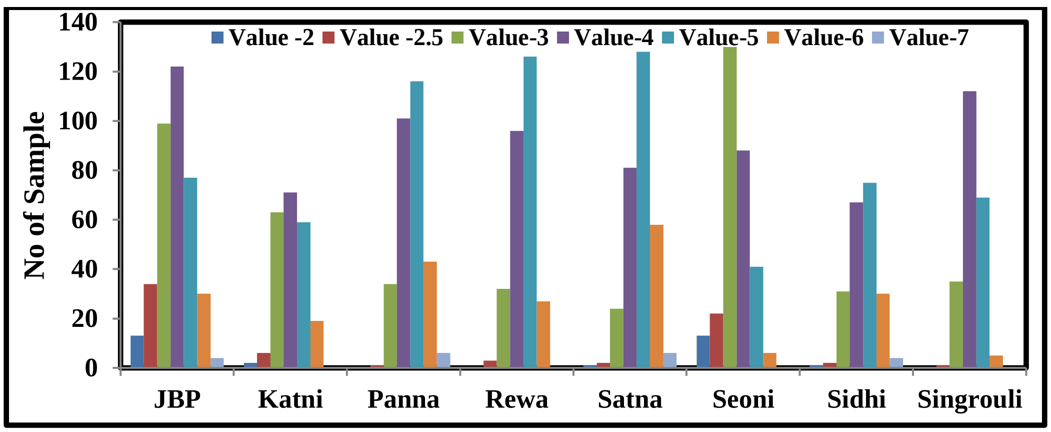



**Fig. 4 Frequency distribution of Value in Kymore plateau and Satpura hills zone**

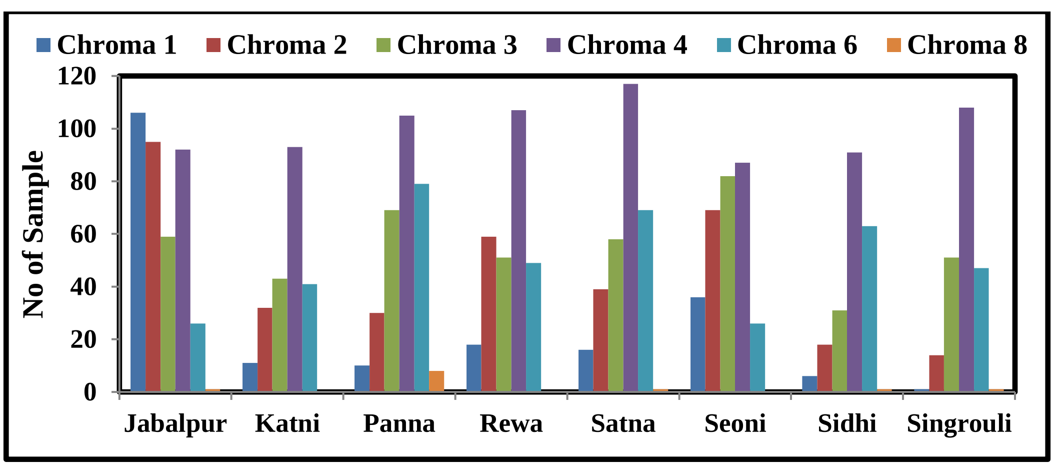



**Fig. 5 Frequency distribution of Chroma in Kymore plateau and Satpura hills zone**

### 3.2 Status of physico-chemical properties under different colour parameters

The study of soil properties such as pH, EC, and SOC across different Munsell colour parameters (Hue, Value, and Chroma) revealed notable variations in soil chemistry across the different districts

#### 3.2.1pH, EC and SOC content in soils under different hue

Data presented in Table 2 showed that the pH ranged from 5.82 to 7.30, 4.86 to 8.17, 4.79 to 8.22, 5.00 to 8.02, 6.30 to 7.17, 5.56 to 8.10, 5.24 to 7.92 and 6.55 to 8.39 with a mean value of 6.57, 6.64, 6.82, 6.53, 6.78, 7.02, 6.53 and 6.61at hue of 10R, 10YR, 2.5Y, 2.5YR, 5R, 5Y, 5YR and 7.5YR, respectively. The EC ranged from 0.14 to 0.29, 0.05 to 0.96,0.03 to 0.97, 0.05 to 0.54, 0.19 to .22, 0.08 to 0.66, 0.08 to 0.73 and 0.04 to 0.97 dS/m with a mean value 0.20, 0.25, 0.26, 0.21, 0.20, 0.25, 0.24 and 0.26 dS/m at hue of 10R, 10YR, 2.5Y, 2.5YR, 5R, 5Y, 5YR and 7.5YR, respectively. However, the SOC ranged from 4.70 to 6.98, 1.50 to 12.45,1.50 to 10.95, 1.20 to 9.75,4.73 to 6.03, 1.50 to 7.05, 2.10 to 11.70 and 1.05 to 10.65 g/kg with a mean value of 5.56, 5.48, 5.64, 5.27, 5.55, 4.77, 5.43 and 5.35 g/kg at hue of 10R, 10YR, 2.5Y, 2.5YR, 5R, 5Y, 5YR and 7.5YR, respectively.

Soils with hues of 10YR and 5Y exhibited relatively higher pH levels, suggesting slightly alkaline conditions, while soils with hues like 2.5Y and 2.5YR showed a more neutral to slightly acidic pH range. EC values were low to moderate, ranging from 0.03 to 0.97 dS/m, with mean values ranging between 0.20 and 0.26 dS/m. This indicates that the soils generally have low salinity. SOC content showed a wide range, from 1.05 to 12.45 g/kg, with mean values varying between 4.77 and 5.64 g/kg. Soils with hues such as 10YR and 2.5Y generally had higher SOC content, reflecting better organic matter accumulation. These findings align with studies by Brady and Weil (2008), which suggest that pH and EC are influenced by the mineral composition and organic matter content of the soil. Additionally, soils with higher SOC typically show better fertility and moisture retention capabilities.

#### 3.2.2 pH, EC and SOC content in soils under different value

The result on pH EC and SOC under different value was presented in Table 2 showed that the pH ranged from 5.95 to 7.81, 5.25 to 8.10, 5.00 to 8.10, 4.55 to 8.39, 4.79 to 8.22, 4.87 to 8.11 and 5.41 with a mean value of 6.94, 6.96, 6.70, 6.68, 6.72, 6.75 and 6.91 at value of 2, 2.5, 3, 4, 5, 6 and 7, respectively. The EC ranged from0.09 to 0.92, 0.09 to 0.58, 0.05 to 0.85, 0.03 to 0.97, 0.05 to 0.97, 0.05 to 0.91 and 0.07 to 0.53 dS/m with a mean value of 0.24, 0.26, 0.24, 0.25, 0.26, 0.25 and 0.26dS/m at value of 2, 2.5, 3, 4, 5, 6 and 7, respectively. The SOC ranged from 2.70 to 11.25, 1.95 to 10.65,1.50 to 12.45, 1.05 to 10.9, 1.05 to 10.95, 2.25 to 10.35 and 4.42 to 9.30 g/kg with a mean value of 6.20, 5.89,5.66, 4.49, 5.40 5.47 and 5.71 g/kg at value of 2, 2.5, 3, 4, 5, 6 and 7, respectively. The pH values in soils across different value categories ranged from 4.55 to 8.39, with mean values from 6.68 to 6.96, suggesting that soil pH was generally neutral to slightly alkaline. EC ranged from 0.03 to 0.97 dS/m, with mean values ranging between 0.24 and 0.26 dS/m, consistent with low salinity levels in the soils. The SOC content across the different value categories ranged from 1.05 to 12.45 g/kg, with mean values ranging from 4.49 to 6.20 g/kg. The higher SOC content was found in soils with lower value (more intense colour), supporting the idea that darker soils may have more organic matter (Schmidt et al., 2011).

### 3.2.3 pH, EC and SOC content in soils under different chroma

Result from the Table 2 showed that the pH ranged from4.95 to 8.10, 4.79 to 8.22, 4.55 to 8.15, 4.94 to 8.12 and 5.79 to 7.695.14 to 8.39 with a mean value of 7.00, 7.85, 6.73, 6.62, 6.63 and 6.93 at chroma of 1, 2, 3, 4, 6 and 8, respectively. The EC ranged from 0.03 to 0.92, 0.05 to 0.73, 0.07 to 0.82, 0.04 to 0.97, 0.05 to 0.96 and 0.09 to 0.43 dS/m with a mean value of 0.25, 0.26, 0.25,0.25, 0.25 and 0.28dS/m at chroma of 1, 2, 3, 4, 6 and 8, respectively. The OC ranged from 1.50 to 12.45, 1.50 to 10.95, 1.65 to 10.80, 1.05 to 11.70, 1.05 to 9.30 and 3.60 to 7.95g/kg with a mean value of 5.84, 5.79, 5.53,5.44, 5.26 and 5.36 g/kg at chroma of 1, 2, 3, 4, 6 and 8, respectively. Soil pH varied from 4.55 to 8.39 across different chroma values, with mean pH values ranging from 6.62 to 7.85, indicating slight alkalinity. The EC values ranged from 0.03 to 0.97 dS/m, with mean values ranging from 0.25 to 0.28 dS/m, suggesting that the soils were mostly non-saline. SOC content ranged from 1.05 to 12.45 g/kg, with mean values from 5.26 to 5.84 g/kg. Soils with higher chroma (more vivid colour) had slightly lower SOC values, indicating that chroma might be less correlated with OC content but could be influenced by factors such as soil texture and moisture.

**Table 2: Physico-chemical parameter of soils under different Munsell colour parameters**

| Parameter | | n | pH | | | | EC ($dSm^{-1}$) | | | | OC ($g\ kg^{-1}$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Min | Max | Mean | CV (%) | Min | Max | Mean | CV (%) | Min | Max | Mean | CV (%) |
| Hue | 10R | 5 | 5.82 | 7.30 | 6.57 | 8.83 | 0.14 | 0.29 | 0.20 | 35.46 | 4.70 | 6.98 | 5.56 | 15.73 |
| | 10YR | 611 | 4.86 | 8.17 | 6.64 | 10.27 | 0.05 | 0.96 | 0.25 | 49.52 | 1.50 | 12.45 | 5.48 | 23.49 |
| | 2.5Y | 1074 | 4.79 | 8.22 | 6.82 | 9.20 | 0.03 | 0.97 | 0.26 | 49.82 | 1.50 | 10.95 | 5.64 | 22.86 |
| | 2.5YR | 113 | 5.00 | 8.02 | 6.53 | 9.61 | 0.05 | 0.54 | 0.21 | 45.68 | 1.20 | 9.75 | 5.27 | 23.01 |
| | 5R | 3 | 6.30 | 7.17 | 6.78 | 6.52 | 0.19 | 0.22 | 0.20 | 7.51 | 4.73 | 6.03 | 5.55 | 12.86 |
| | 5Y | 33 | 5.36 | 8.10 | 7.02 | 8.91 | 0.08 | 0.66 | 0.25 | 51.76 | 1.50 | 7.05 | 4.77 | 24.10 |
| | 5YR | 66 | 5.24 | 7.92 | 6.53 | 9.49 | 0.08 | 0.73 | 0.24 | 58.25 | 2.10 | 11.70 | 5.43 | 28.70 |
| | 7.5YR | 311 | 4.55 | 8.39 | 6.61 | 9.95 | 0.04 | 0.97 | 0.26 | 58.07 | 1.05 | 10.65 | 5.35 | 20.93 |
| Value | 2 | 30 | 5.95 | 7.81 | 6.94 | 7.13 | 0.09 | 0.92 | 0.24 | 74.12 | 2.70 | 11.25 | 6.20 | 26.18 |
| | 2.5 | 71 | 5.25 | 8.10 | 6.96 | 8.70 | 0.09 | 0.58 | 0.26 | 41.37 | 1.95 | 10.65 | 5.89 | 26.94 |
| | 3 | 448 | 5.00 | 8.10 | 6.70 | 9.35 | 0.05 | 0.85 | 0.24 | 49.99 | 1.50 | 12.45 | 5.66 | 23.67 |
| | 4 | 738 | 4.55 | 8.39 | 6.68 | 9.79 | 0.03 | 0.97 | 0.25 | 51.72 | 1.05 | 10.95 | 5.49 | 23.16 |
| | 5 | 691 | 4.79 | 8.22 | 6.72 | 9.68 | 0.05 | 0.97 | 0.26 | 51.15 | 1.05 | 10.95 | 5.40 | 22.80 |
| | 6 | 218 | 4.87 | 8.11 | 6.75 | 11.02 | 0.05 | 0.91 | 0.25 | 52.99 | 2.25 | 10.35 | 5.47 | 19.10 |
| | 7 | 20 | 5.41 | 7.82 | 6.91 | 9.06 | 0.07 | 0.53 | 0.26 | 48.48 | 4.42 | 9.30 | 5.71 | 22.85 |
| Chroma | 1 | 204 | 4.95 | 8.10 | 7.00 | 8.37 | 0.03 | 0.92 | 0.25 | 50.86 | 1.50 | 12.45 | 5.84 | 27.58 |
| | 2 | 356 | 4.79 | 8.22 | 6.85 | 9.36 | 0.05 | 0.73 | 0.26 | 47.20 | 1.50 | 10.95 | 5.79 | 24.34 |
| | 3 | 444 | 5.14 | 8.39 | 6.73 | 9.61 | 0.07 | 0.82 | 0.25 | 47.96 | 1.65 | 10.80 | 5.53 | 22.43 |
| | 4 | 800 | 4.55 | 8.15 | 6.62 | 9.89 | 0.04 | 0.97 | 0.25 | 54.75 | 1.05 | 11.70 | 5.44 | 21.98 |
| | 6 | 400 | 4.94 | 8.12 | 6.63 | 9.87 | 0.05 | 0.96 | 0.25 | 52.49 | 1.05 | 9.30 | 5.26 | 20.41 |
| | 8 | 12 | 5.79 | 7.69 | 6.93 | 8.35 | 0.09 | 0.43 | 0.28 | 36.17 | 3.60 | 7.95 | 5.36 | 19.15 |

### 3.3 Status of macro nutrients in soils under different colour parameters

The availability of P, K and S varied significantly in soils based on their Munsell colour parameters, including Hue, Value, and Chroma.

#### 3.3.1 Availability of P, K and S in soils under different Hue

Data given in Table 3 revealed that the available P ranged from 13.59 to 40.79, 1.44 to 74.02, 1.15 to 117.17, 1.19 to 74.02, 5.17 to 7.77, 3.43 to 29.22, 1.73 to 26.32 and 1.11 to 67.12 kg $ha^{-1}$with a mean value of 20.65, 10.22, 11.20, 10.59, 6.32, 9.97, 8.74 and 9.86 kg $ha^{-1}$at hue of 10R, 10YR, 2.5Y, 2.5YR, 5R, 5Y, 5YR and 7.5YR, respectively. The available K ranged from 290.08 to 386.40, 116.48 to 715.68, 105.28 to 974.40, 122.08 to 611.52,244.16 to 324.80, 126.56 to 557.76, 147.84 to 611.52 and 145.60 to 750.40 kg $ha^{-1}$with a mean value of 343, 344, 366, 340, 293, 355, 324 and 327 kg $ha^{-1}$at hue of 10R, 10YR, 2.5Y, 2.5YR, 5R, 5Y, 5YR and 7.5YR, respectively. The available S ranged from 3.12 to 8.79, 0.43 to 33.60, 0.57 to 30.77, 0.57 to 19.00, 5.10 to 7.66, 2.55 to 19.99, 0.99 to 20.84 and 0.71 to 18.29 $mgkg^{-1}$with a mean value of

7.34, 7.71, 8.21, 6.95, 6.57, 9.04, 6.82 and 6.78 mg kg$^{-1}$at hue of 10R, 10YR, 2.5Y, 2.5YR, 5R, 5Y, 5YR and 7.5YR, respectively. The available P in soils varied widely across different hues, ranging from 1.11 to 117.17 kg/ha, with mean values varying between 6.32 and 20.65 kg/ha. The highest P availability was observed in soils with hues of 2.5Y, 10R, and 2.5YR. The K availability ranged from 105.28 to 974.40 kg/ha, with mean values from 293 to 366 kg/ha. Soils with hues like 10YR and 5Y showed the highest potassium levels. The S availability ranged from 0.43 to 33.60 mg/kg, with mean values between 6.57 and 9.04 mg/kg. Soils with hues such as 10YR and 5Y exhibited the highest sulfur availability. These results align with findings by Conklin et al. (2013) who highlighted that soil colour can be an indirect indicator of nutrient content due to the relationships between soil organic matter, mineralogy, and nutrient availability.

### 3.3.2 Availability of P, K and S in soils under different Value

From the result presented in Table 3 showed the available P ranged from 1.73 to 50.05, 1.44 to 67.12, 1.44 to 61.62, 1.19 to 96.30, 1.11 to 117.7, 1.44 to 74.02 and 2.89 to 28.64 kg ha$^{-1}$ with a mean value of 15.32, 13.56, 10.35, 10.23, 10.61, 11.05 and 10.83 kg ha$^{-1}$ at value of 2, 2.5, 3, 4, 5, 6 and 7, respectively. The available K ranged from 136.64 to 624.96, 178.08 to 750.40,116.48 to 678.72, 125.44 to 738.08, 122.08 to 848.96, 160.16 to 974.40 and 105.28 to 468.16 kg ha$^{-1}$ with a mean value of376, 388, 355, 349, 348, 354 and 301 kg ha$^{-1}$ at value of 2, 2.5, 3, 4, 5, 6 and 7, respectively. The available S ranged from 2.13 to 33.60, 2.55 to 30.77, 0.57 to 29.77, 0.43 to 30.34, 0.57 to 26.23, 0.85 to 27.51 and 2.41 to 18.57 mg kg$^{-1}$with a mean value of 8.69, 7.96, 7.66, 7.84, 7.62, 8.03 and 8.02 mg kg$^{-1}$at value of 2, 2.5, 3, 4, 5, 6 and 7, respectively.

The P ranged from 1.11 to 117.7 kg/ha, with mean values between 10.23 and 15.32 kg/ha. Higher P levels were found in soils with lower value (darker soil colours). The K availability ranged from 105.28 to 848.96 kg/ha, with a mean value from 301 to 388 kg/ha. The S availability varied from 0.43 to 33.60 mg/kg, with mean values ranging from 7.62 to 8.69 mg/kg. The increase in phosphorus and sulfur content in soils with lower values (more intense colours) could reflect higher organic matter content, which can influence nutrient release (Wiesmeier et al., 2019).

### 3.3.3Availability of P, K and S in soils under different chroma

Data on available P, K and S in Table 3 indicated that the available P ranged from 1.44 to 96.30, 1.44 to 95.18, 1.11 to 74.02, 1.15 to 55.84, 1.19 to 117.17 and 2.02 to 11.28 kg ha$^{-1}$with a mean value of13.73, 12.32, 10.22, 10.05, 9.30 and 6.94 kg ha$^{-1}$at chroma of1, 2, 3, 4, 6 and 8, respectively. The available K ranged from 136.64 to 848.96, 126.56 to 750.40, 116.48 to 721.28, 114.24 to 743.68, 105.28to 974.40 and 252.00 to 543.20kg ha$^{-1}$with a mean value of 397, 376, 348, 341, 329 and 389 kg ha$^{-1}$atchroma of1, 2, 3, 4, 6 and 8, respectively. The available S ranged from 1.13 to 33.60, 0.71 to 30.34,0.85 to 29.77 0.43 to 25.38, 0.57 to 23.25 and 2.98 to 22.12 mg kg$^{-1}$with a mean value of 8.69, 8.51, 7.82,7.54, 7.07 and 7.40 mg kg$^{-1}$at chroma of 1, 2, 3, 4, 6 and 8, respectively. The availability of P ranged from 1.11 to 117.17 kg/ha, with mean values varying between 6.94 and 13.73 kg/ha.

The highest P availability was found in soils with higher chroma (more vivid colours), particularly those with chroma of 1 and 2. Potassium (K) availability ranged from 105.28 to 974.40 kg/ha, with a mean value between 329 and 397 kg/ha. Sulfur (S) availability varied from 0.43 to 33.60 mg/kg, with mean values between 7.07 and 8.69 mg/kg. Higher chroma soils exhibited relatively higher K and S levels, possibly due to differences in soil texture and mineral composition (Ibáñez-Asensio et al., 2013). The availability of P, K, and S is strongly influenced by soil colour parameters, soils with darker hues (lower value) and higher chroma tend to have higher nutrient availability, likely due to better organic matter content, which enhances nutrient retention and release. The high variability in phosphorus and sulfur availability in soils with different hues, values, and chromas underscores the need for targeted soil management strategies based on soil colour to optimize nutrient availability.

**Table 3: Macro-nutrients status under different Munsell colour parameters**

| Parameter | | n | P (kg ha$^{-1}$) | | | | K (kg ha$^{-1}$) | | | | S (mg kg$^{-1}$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Min | Max | Mean | CV (%) | Min | Max | Mean | CV (%) | Min | Max | Mean | CV (%) |
| Hue | 10R | 5 | 13.59 | 40.79 | 20.65 | 56.22 | 290 | 386 | 343 | 11.10 | 3.12 | 8.79 | 7.34 | 32.49 |
| | 10YR | 611 | 1.44 | 74.02 | 10.22 | 75.02 | 116 | 716 | 344 | 28.73 | 0.43 | 33.60 | 7.71 | 48.35 |
| | 2.5Y | 1074 | 1.15 | 117.17 | 11.20 | 85.06 | 105 | 974 | 366 | 29.89 | 0.57 | 30.77 | 8.21 | 48.71 |
| | 2.5YR | 113 | 1.19 | 74.02 | 10.59 | 86.10 | 122 | 612 | 340 | 29.79 | 0.57 | 19.00 | 6.95 | 45.25 |
| | 5R | 3 | 5.17 | 7.77 | 6.32 | 20.97 | 244 | 325 | 293 | 14.72 | 5.10 | 7.66 | 6.57 | 20.05 |
| | 5Y | 33 | 3.43 | 29.22 | 9.97 | 57.03 | 127 | 558 | 355 | 29.64 | 2.55 | 19.99 | 9.04 | 44.90 |
| | 5YR | 66 | 1.73 | 26.32 | 8.74 | 63.87 | 148 | 612 | 324 | 30.17 | 0.99 | 20.84 | 6.82 | 43.32 |
| | 7.5YR | 311 | 1.11 | 67.12 | 9.86 | 74.13 | 146 | 750 | 327 | 26.77 | 0.71 | 18.29 | 6.78 | 44.80 |
| Value | 2 | 30 | 1.73 | 50.05 | 15.32 | 81.44 | 137 | 625 | 376 | 34.22 | 2.13 | 33.60 | 8.69 | 71.93 |
| | 2.5 | 71 | 1.44 | 67.12 | 13.56 | 86.10 | 178 | 750 | 388 | 30.43 | 2.55 | 30.77 | 7.96 | 49.82 |
| | 3 | 448 | 1.44 | 61.62 | 10.35 | 65.95 | 116 | 679 | 355 | 27.46 | 0.57 | 29.77 | 7.66 | 45.45 |
| | 4 | 738 | 1.19 | 96.30 | 10.23 | 81.50 | 125 | 738 | 349 | 30.41 | 0.43 | 30.34 | 7.84 | 50.37 |
| | 5 | 691 | 1.11 | 117.17 | 10.61 | 84.10 | 122 | 849 | 348 | 28.68 | 0.57 | 26.23 | 7.62 | 46.02 |
| | 6 | 218 | 1.44 | 74.02 | 11.05 | 88.82 | 160 | 974 | 354 | 30.90 | 0.85 | 27.51 | 8.03 | 48.82 |
| | 7 | 20 | 2.89 | 28.64 | 10.83 | 59.63 | 105 | 468 | 301 | 30.61 | 2.41 | 18.57 | 8.02 | 55.19 |
| Chroma | 1 | 204 | 1.44 | 96.30 | 13.73 | 91.37 | 137 | 849 | 397 | 30.13 | 1.13 | 33.60 | 8.69 | 54.51 |
| | 2 | 356 | 1.44 | 95.18 | 12.32 | 82.47 | 127 | 750 | 376 | 29.37 | 0.71 | 30.34 | 8.51 | 47.97 |
| | 3 | 444 | 1.11 | 74.02 | 10.22 | 76.45 | 116 | 721 | 348 | 29.64 | 0.85 | 29.77 | 7.82 | 46.52 |
| | 4 | 800 | 1.15 | 55.84 | 10.05 | 69.30 | 114 | 744 | 341 | 28.43 | 0.43 | 25.38 | 7.54 | 45.45 |
| | 6 | 400 | 1.19 | 117.17 | 9.30 | 85.99 | 105 | 974 | 329 | 27.56 | 0.57 | 23.25 | 7.07 | 49.14 |
| | 8 | 12 | 2.02 | 11.28 | 6.94 | 47.45 | 252 | 543 | 389 | 24.15 | 2.98 | 22.12 | 7.40 | 67.05 |

### 3.4 Correlation findings

The correlation results between soil colour parameters (CIE Lab and Munsell) and various soil properties. These correlations provide insights into how colour characteristics can be used as indicators for soil fertility and nutrient availability. Correlation result showed in Table 4 indicated that the CIE lab parameters *i.e.,* L* was significantly correlated with SOC (r=-0.088**) and K (r= -0.062**). Further, a* and b* parameters were found significantly correlated with all the parameters except EC which was showed non-significant correlation with b*.. The SOC was significantly related with R, G and B with r value of -0.114**, -0.071**) and-0.051*) but EC showed non-significant correlation with R, G and B colour component, respectively. Munsell parameters i.e., Value showed significant relation with OC (r=-0.82**) and available K (r=-0.51*). This finding is in line with previous research indicating that darker soils often contain higher amounts of organic carbon and nutrients (Bohn et al., 2001). In contrast, Chroma was chroma was negatively correlated with OC (r=-0.141**), available P (r=-0.143**), K (r=-0.173**) and S (r=-0.134**). Soils with higher chroma (more vivid colour) tended to have lower nutrient availability, which could be due to higher soil mineralization and less organic matter retention in these soils Minh *et al.,* (2023). Djama *et al.,* (2023) and Kang *et al.,* (2024) also found negatively significant correlation between SOC and soil colour component (RGB). The correlations observed between colour parameters and soil properties confirm the utility of soil colour as an indicator of fertility. Lighter soils (higher L* values) tend to have lower organic content and nutrient availability, while darker soils (lower value) are generally more fertile, with higher organic carbon and potassium content. The negative correlation of chroma with several soil nutrients suggests that soil colour, particularly in terms of chroma and value, can be a helpful tool for predicting soil fertility and guiding soil management practices.

**Table 4: Correlation between soil colour and soil chemical properties**

| Parameters | Munsell parameters | | CIE-Lab colour space | | | RGB parameters | | |
|---|---|---|---|---|---|---|---|---|
| | Value | Chroma | L* | a* | b* | Red | Green | Blue |
| **pH** | -0.002 | -0.149** | -0.015 | -0.160** | -0.120** | -0.065** | 0.004 | 0.056** |
| **EC** | 0.028 | -0.013 | 0.019 | -0.058** | -0.016 | 0.002 | 0.029 | 0.017 |
| **OC** | -0.082** | -0.141** | -0.088** | -0.115** | -0.105** | -0.114** | -0.071** | -0.101** |
| **Avail. P** | -0.015 | -0.143** | -0.019 | -0.095** | -0.106** | -0.053* | -0.002 | 0.055* |
| **Avail. K** | -0.051* | -0.173** | -0.062** | -0.148** | -0.117** | -0.098** | -0.045* | -0.003 |
| **Avail.S** | -0.000 | -0.134** | -0.001 | -0.158** | -0.089** | -0.042 | 0.017 | 0.048* |

*= significant at 5%, **= significant at 1%,

Further, the Fig. 6, 7 and 8 depicted the relationship between soil colour parameter value and chroma. The best fitted with the models were polynomial relationship between the value and Chroma of colour parameter and SOC, equation of $Y=0.086x^2-0.860x+7.528$ with a $R^2$ is 0.982 and $Y=0.018x^2-0.249x+6.126$ with a $R^2$ is 0.948, respectively. However, the linear relationship was found between the chroma and available P, Y=-0.873+13.92 with a $R^2$ is 0.922. This is supported by research by Nair et al., (2009), which suggests that soils with higher chroma often have better P availability due to higher microbial activity and organic matter content. The strong polynomial correlations between soil colour (value and chroma) and SOC emphasize the potential of using soil colour as an indicator of organic matter content. The linear correlation between chroma and available phosphorus further supports the use of colour as a quick diagnostic tool for assessing phosphorus levels in soils.

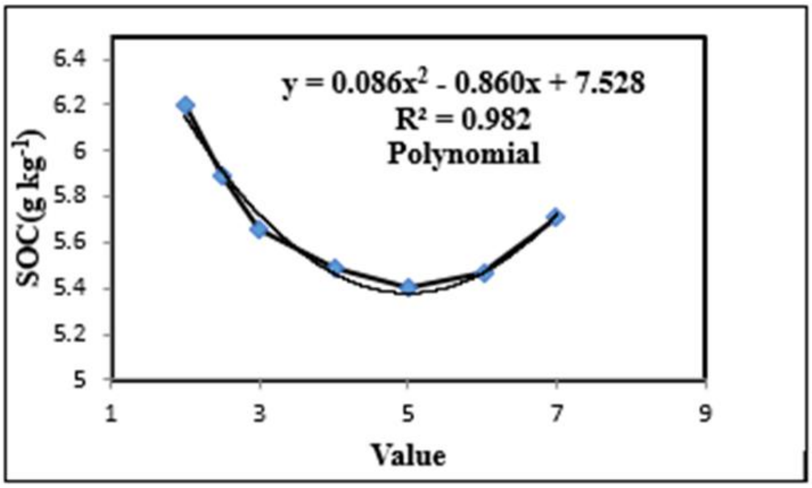



Fig 6. Relationship between Value and SOC

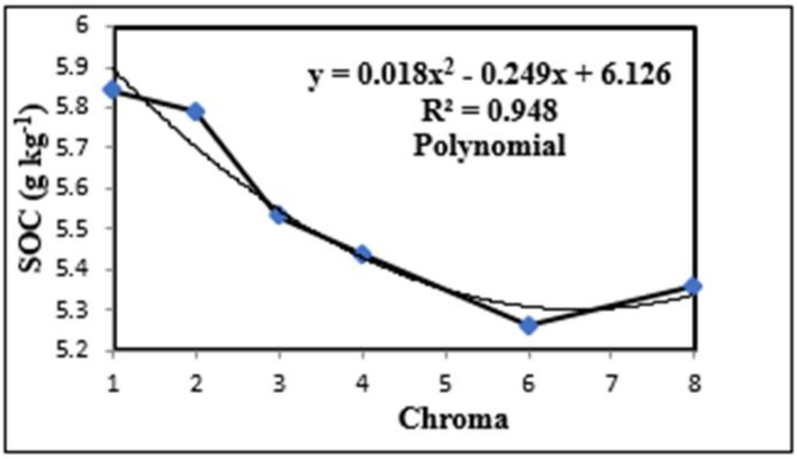



Fig 7. Relationship between Chroma and SOC

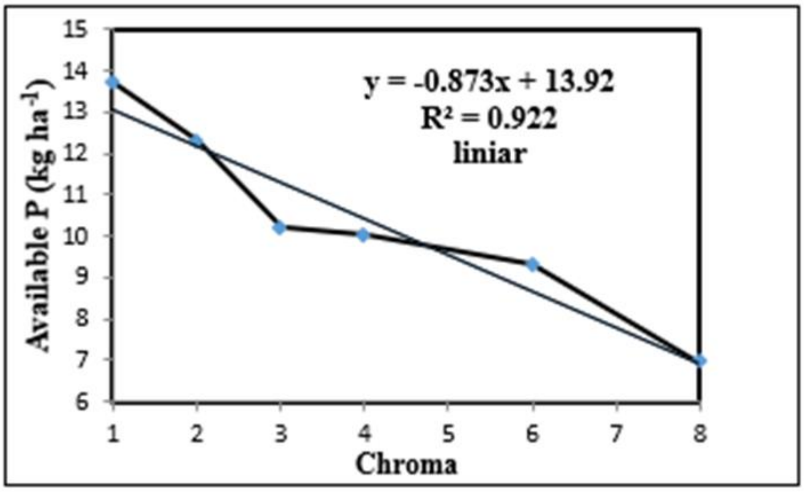



Fig 8. Relationship between Chroma and Available P

### 3.4 Prediction of soil colour parameters using spectroscopic and different machine learning models

The performance of various machine learning models in predicting soil colour parameters (L*, a*, and b*) was evaluated based on the coefficient of determination ($R^2$), root mean square error (RMSE), and residual prediction deviation (RPD) for both the training and testing datasets (Figure 9)

#### 3.5.1 CIE-Lab colour space

The performance results presented in the Table 5 showed the PLSR, SVR, ANN, Random Forest, XGBoost, LightGBM, CatBoost and ELM models showed the $R^2$ of 0.24, 0.18, 0.28, 0.84, 0.84, 0.84, 0.83 and 0.38; 0.32, 0.28, 0.38, 0.81, 0.81, 0.81, 0.81 and 0.31 and 0.20, 0.12, 0.23, 0.83, 0.83, 0.83, 0.82 and 0.20 with a RMSE and RPD of 9.43, 9.79, 9.14, 5.25, 5.28, 5.26, 5.30 and 9.46; 3.99, 4.40, 3.83, 2.37, 2.37, 2.38, 2.35 and 4.04; 8.48, 9.03, 8.32, 4.85, 4.82, 4.84, 4.86 and 8.50 and 1.14, 1.10, 1.18, 2.06, 2.05, 2.05, 2.04 and 1.14; 1.22, 1.10, 1.27, 2.04, 2.05, 2.04, 2.07 and 1.20 and 1.12, 1.05, 1.14, 1.96, 1.97, 1.96, 1.95 and 1.12 for L*, a* and b* for a training dataset. However, PLSR, SVR, ANN, Random Forest, XGBoost, LightGBM, CatBoost and ELM models showed $R^2$ of 0.33, 0.29, 0.36, 0.36, 0.36, 0.35, 0.36 and 0.35; 0.37, 0.33, 0.35, 0.31, 0.31, 0.32, 0.32 and 0.32 and 0.30, 0.22, 0.24, 0.36, 0.37, 0.37, 0.36 and 0.31 with a RMSE and RPD of 7.23, 7.41, 7.11, 7.10, 7.07, 7.17, 7.11 and 7.14; 3.17, 3.41, 3.27, 3.33, 3.33, 3.29, 3.29 and 3.30 and 6.36, 6.72, 6.72, 6.05, 6.03, 6.06, 6.08 and 6.31 and 1.22, 1.19, 1.24, 1.24, 1.25, 1.23, 1.24and1.23; 1.25, 1.16, 1.21, 1.21, 1.19, 1.20, 1.20 and 1.20 and 1.19, 1.12, 1.12, 1.25, 1.25, 1.25, 1.24 and 1.20 for the testing dataset, respectively. Overall the XGBoost, Random Forest and XGBoost model gave the best prediction of soil colour parameter of L*, a* and b*.

For the training dataset, models such as XGBoost, Random Forest, and LightGBM demonstrated the highest $R^2$ values of 0.84, 0.84, and 0.84, respectively, for predicting L*, a*, and b* values. These models exhibited relatively low RMSE and RPD values (around 5.25 to 5.30 and 2.04 to 2.07, respectively), indicating strong predictive performance. In contrast, models like PLSR, SVR, and ANN had lower $R^2$ values (ranging from 0.18 to 0.28) and higher RMSE and RPD values, suggesting poorer model performance. On the testing dataset, XGBoost, Random Forest, and LightGBM again outperformed other models, achieving $R^2$ values of 0.36, 0.36, and 0.35, respectively, across the L*, a*, and b* colour components. These models also had lower RMSE and RPD values (around 7.07 to 7.17 and 1.19 to 1.24, respectively), indicating robust generalization capabilities. The performance of these models was consistent, supporting their suitability for predicting soil colour parameters. Conversely, other models such as PLSR and SVR exhibited lower $R^2$ values (ranging from 0.29 to 0.33) and higher RMSE and RPD values, highlighting their limitations. The superior performance of the XGBoost, Random Forest, and LightGBM models. These models effectively capture the non-linear relationships between soil colour and properties such as organic carbon and nutrient content, which is critical for soil management and precision agriculture. In contrast, traditional models like PLSR and SVR, although simpler, struggle with the complexity of soil data, resulting in lower prediction accuracy.

**Table 5: Results of different model for prediction of soil colour parameter (L*, a* and b*)**

| Property | Model | Training (80%) | | | Testing data (20%) | | |
|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD |
| L* | PLSR | 0.24 | 9.43 | 1.14 | 0.33 | 7.23 | 1.22 |
| | SVR | 0.18 | 9.79 | 1.10 | 0.29 | 7.41 | 1.19 |
| | ANN | 0.28 | 9.14 | 1.18 | 0.36 | 7.11 | 1.24 |
| | RandomForest | 0.84 | 5.25 | 2.06 | 0.36 | 7.10 | 1.24 |
| | **XGBoost** | **0.84** | **5.28** | **2.05** | **0.36** | **7.07** | **1.25** |
| | LightGBM | 0.84 | 5.26 | 2.05 | 0.35 | 7.17 | 1.23 |
| | CatBoost | 0.83 | 5.30 | 2.04 | 0.36 | 7.11 | 1.24 |
| | ELM | 0.23 | 9.46 | 1.14 | 0.35 | 7.14 | 1.23 |
| a* | PLSR | 0.32 | 3.99 | 1.22 | 0.37 | 3.17 | 1.25 |
| | SVR | 0.28 | 4.40 | 1.10 | 0.33 | 3.41 | 1.16 |
| | ANN | 0.38 | 3.83 | 1.27 | 0.35 | 3.27 | 1.21 |
| | **RandomForest** | **0.81** | **2.37** | **2.04** | **0.31** | **3.33** | **1.21** |
| | XGBoost | 0.81 | 2.37 | 2.05 | 0.31 | 3.33 | 1.19 |
| | LightGBM | 0.81 | 2.38 | 2.04 | 0.32 | 3.29 | 1.20 |
| | CatBoost | 0.81 | 2.35 | 2.07 | 0.32 | 3.29 | 1.20 |
| | ELM | 0.31 | 4.04 | 1.20 | 0.32 | 3.30 | 1.20 |
| b* | PLSR | 0.20 | 8.48 | 1.12 | 0.30 | 6.36 | 1.19 |
| | SVR | 0.12 | 9.03 | 1.05 | 0.22 | 6.72 | 1.12 |
| | ANN | 0.23 | 8.32 | 1.14 | 0.24 | 6.72 | 1.12 |
| | RandomForest | 0.83 | 4.85 | 1.96 | 0.36 | 6.05 | 1.25 |
| | **XGBoost** | **0.83** | **4.82** | **1.97** | **0.37** | **6.03** | **1.25** |
| | LightGBM | 0.83 | 4.84 | 1.96 | 0.37 | 6.06 | 1.25 |
| | CatBoost | 0.82 | 4.86 | 1.95 | 0.36 | 6.08 | 1.24 |
| | ELM | 0.20 | 8.50 | 1.12 | 0.31 | 6.31 | 1.20 |

### 3.5.2 Spectral colour-R, G, B

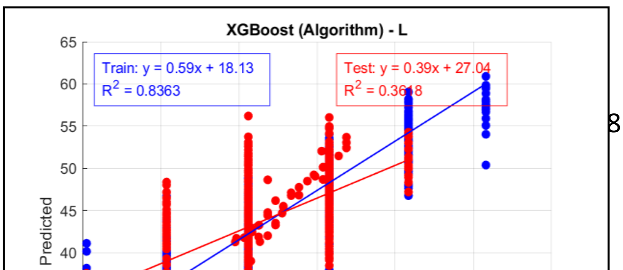
The performance results presented in the Table 6 indicated that the PLSR, SVR, ANN, Random Forest, XGBoost, LightGBM, CatBoost and ELM models showed the $R^2$ of 0.25, 0.18, 0.37, 0.84, 0.85, 0.84, 0.84 and 0.0.25; 0.23, 0.16, 0.29, 0.84, 0.84, 0.84, 0.84 and 0.22 and 0.16, 0.08, 0.16, 0.84, 0.83, 0.83, 0.84 and 0.15 with a RMSE and RPD of 25.60, 26.77, 23.60, 14.05, 14.01, 14.12, 14.04 and 25.67; 22.18, 23.22, 21.43, 12.41, 12.32, 12.38, 12.40 and 22.30; 19.96, 20.93, 20.08, 11.38, 11.37, 11.38, 11.34 and 20.12 and 1.15, 1.10, 1.25, 2.10, 2.11, 2.09, 2.11 and 1.15; 1.14, 1.09, 1.18, 2.04, 2.06, 2.05, 2.04 and 1.14 and 1.09, 1.04, 1.09, 1.92, 1.92, 1.92, 1.92 and 1.08 for R, G and B for a training dataset. However, PLSR, SVR, ANN, Random Forest, XGBoost, LightGBM, CatBoost and ELM models showed $R^2$ of 0.37, 0.37, 0.31, 0.45, 0.45, 0.45, 0.45 and 0.38; 0.39, 0.34, 0.41, 0.43, 0.43, 0.43, 0.43 and 0.42 and 0.23, 0.16, 0.09, 0.30, 0.31, 0.29, 0.30 and 0.22 with a RMSE and RPD of 17.89, 17.83, 19.96, 16.74, 16.68, 16.72, 16.78 and 17.69; 15.23, 15.68, 15.25, 14.70, 14.71, 14.74, 14.68 and 14.85 and 13.61, 14.06, 16.06, 12.98, 12.88, 13.06, 13.03 and 13.70 and 1.25, 1.26, 1.12, 1.34, 1.35, 1.34, 1.34 and1.27; 1.27, 1.24, 1.27, 1.32, 1.32, 1.31, 1.32 and 1.30 and 1.13, 1.09, 0.95, 1.18, 1.19, 1.17, 1.18 and 1.12 for the testing dataset, respectively.

In the training dataset, XGBoost, Random Forest, and LightGBM models demonstrated the highest $R^2$ values for predicting the R, G, and B parameters, ranging from 0.83 to 0.85, with RMSE values between 14.01 and 14.12 and RPD values of approximately 2.04 to 2.11. These results indicate a strong predictive ability of these models for soil colour parameters. On the other hand, simpler models like PLSR, SVR, and ANN showed lower $R^2$ values (0.18 to 0.37) and higher RMSE values, suggesting less accurate predictions. For the testing dataset, Random Forest, XGBoost, and LightGBM continued to perform well, achieving $R^2$ values of 0.45 across all RGB components, with RMSE values ranging from 16.68 to 16.78 and RPD values between 1.30 and 1.35. These models exhibited low RMSE and relatively high RPD, demonstrating their robustness and accuracy in predicting soil colour parameters. In contrast, models like PLSR and SVR again showed lower $R^2$ values (ranging from 0.16 to 0.37), with higher RMSE and lower RPD, indicating a weaker performance.

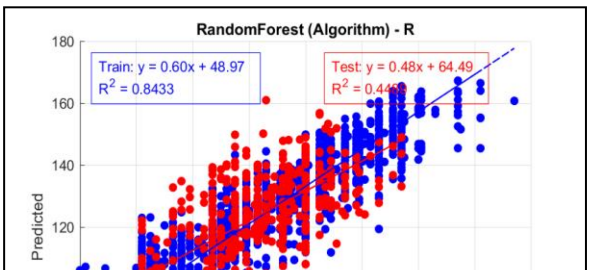
The consistent outperformance of Random Forest, XGBoost, and LightGBM models aligns with recent studies highlighting the effectiveness of ensemble learning techniques in predicting soil properties, including soil colour (Silvero et al. 2021 and Pegalajaret al. 2020). These models can capture complex non-linear relationships between soil characteristics and spectral data, making them ideal for predicting soil colour. In contrast, simpler models like PLSR and SVR struggle with the non-linearities inherent in the data, resulting in poorer prediction accuracy**.** Overall, the Random Forest, Random Forest and XGBoost model gave the best prediction of soil colour parameter of R, G and B. Al-Najiet al. (2022), Ramos et al. (2020), and de Souzaet al. (2022) also used machine learning algorithms for rapid assessment and prediction of soil colour parameters and found some similar trends and supporting finding.

**Table 6: Results of different model for prediction of soil colour parameter (R, G and B)**

| Property | Model | Training data(80%) | | | Testing data (20%) | | |
|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD |
| R | PLSR | 0.25 | 25.60 | 1.15 | 0.37 | 17.89 | 1.25 |
| | SVR | 0.18 | 26.77 | 1.10 | 0.37 | 17.83 | 1.26 |
| | ANN | 0.37 | 23.60 | 1.25 | 0.31 | 19.96 | 1.12 |
| | **RandomForest** | **0.84** | **14.05** | **2.10** | **0.45** | **16.74** | **1.34** |
| | XGBoost | 0.85 | 14.01 | 2.11 | 0.45 | 16.68 | 1.35 |
| | LightGBM | 0.84 | 14.12 | 2.09 | 0.45 | 16.72 | 1.34 |
| | CatBoost | 0.84 | 14.04 | 2.11 | 0.45 | 16.78 | 1.34 |
| | ELM | 0.25 | 25.67 | 1.15 | 0.38 | 17.69 | 1.27 |
| G | PLSR | 0.23 | 22.18 | 1.14 | 0.39 | 15.23 | 1.27 |
| | SVR | 0.16 | 23.22 | 1.09 | 0.34 | 15.68 | 1.24 |
| | ANN | 0.29 | 21.43 | 1.18 | 0.41 | 15.25 | 1.27 |
| | **RandomForest** | **0.84** | **12.41** | **2.04** | **0.43** | **14.70** | **1.32** |
| | XGBoost | 0.84 | 12.32 | 2.06 | 0.43 | 14.71 | 1.32 |
| | LightGBM | 0.84 | 12.38 | 2.05 | 0.43 | 14.74 | 1.31 |
| | CatBoost | 0.84 | 12.40 | 2.04 | 0.43 | 14.68 | 1.32 |
| | ELM | 0.22 | 22.30 | 1.14 | 0.42 | 14.85 | 1.30 |
| B | PLSR | 0.16 | 19.96 | 1.09 | 0.23 | 13.61 | 1.13 |
| | SVR | 0.08 | 20.93 | 1.04 | 0.16 | 14.06 | 1.09 |
| | ANN | 0.16 | 20.08 | 1.09 | 0.09 | 16.06 | 0.95 |
| | RandomForest | 0.84 | 11.38 | 1.92 | 0.30 | 12.98 | 1.18 |
| | **XGBoost** | **0.83** | **11.37** | **1.92** | **0.31** | **12.88** | **1.19** |
| | LightGBM | 0.83 | 11.38 | 1.92 | 0.29 | 13.06 | 1.17 |
| | CatBoost | 0.84 | 11.34 | 1.92 | 0.30 | 13.03 | 1.18 |
| | ELM | 0.15 | 20.12 | 1.08 | 0.22 | 13.70 | 1.12 |





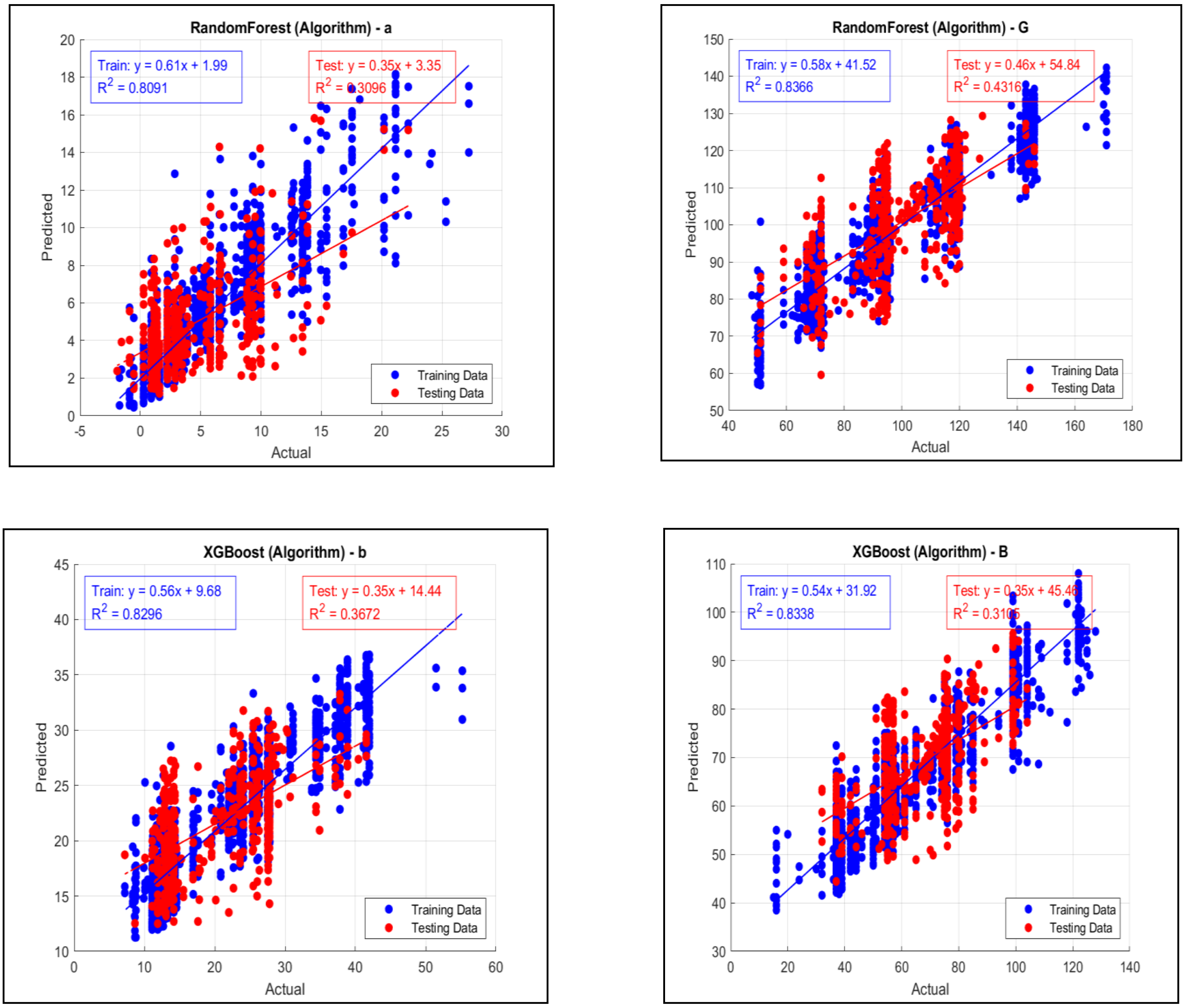
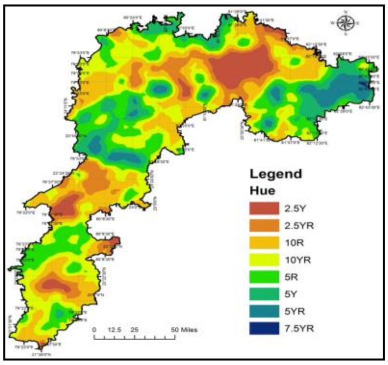


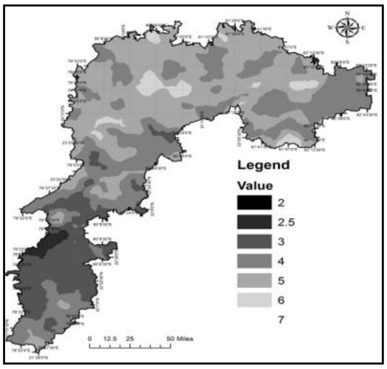

**Fig. 9 Prediction performance of colour parameters (L, a, b, R, G and B)**

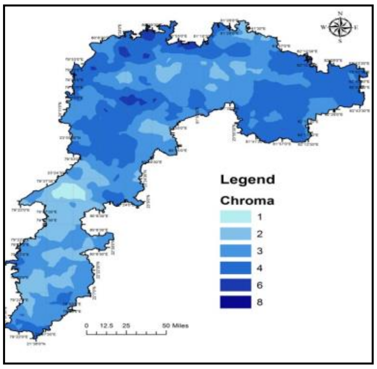
## 3.5 GIS based mapping of soil colour parameters

The spatial variability maps for hue, value, and chroma (Munsell colour parameters) are presented in Fig. 10. The 2.5Y hue was predominantly distributed in the northern region, while 5YR and 7.5YR hues were more variable in the southern part. The value of soils showed higher concentrations in the northern region and lower concentrations in the southern part. The chroma was found to be lower in the southwestern soils of the Kymore Plateau and Satpura Hill zone of Madhya Pradesh, indicating a less vivid soil colour, possibly related to soil texture and mineral composition. These findings are consistent with studies that have linked soil colour variability to the presence of organic matter and minerals, such as manganese oxides, which influence soil redox conditions and consequently the soil's hue and value (Rizzo *et al.*, 2023). The spatial patterns of chroma in these regions also align with findings by other researchers who noted a relationship between chroma and soil mineral content.

**Fig. 10 Spatial variability map of soil colour parameter (a,b,c & d)**

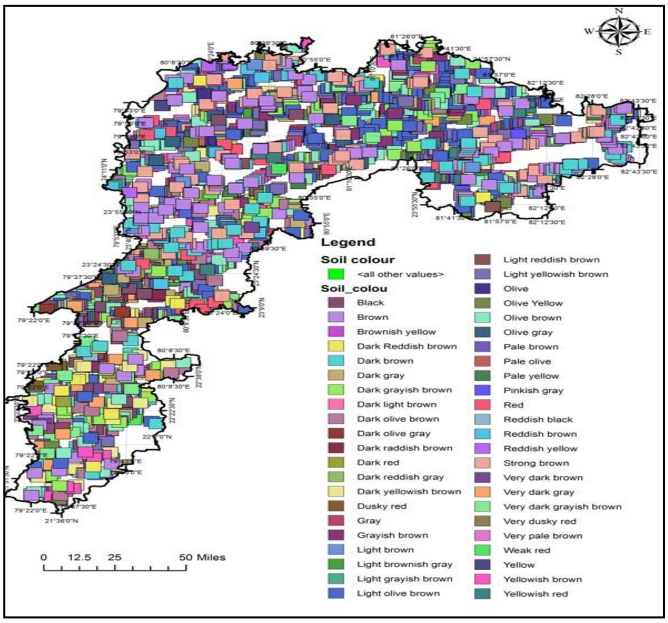



**(a) Hue** **(b) Value** **(c) Chroma**

**(d) Soil colour types**



## 4.0 Conclusions

The study explored the variability in pH, EC, and SOC content across different soil colour parameters (Hue, Value, and Chroma) in the Kymore Plateau and Satpura Hill zones of Madhya Pradesh. It found that soils were generally neutral to slightly alkaline, with low salinity, and substantial variation in organic carbon content. Darker soils (higher value) tended to have more organic carbon. The study highlighted strong correlations between CIE Lab and Munsell colour parameters and soil properties, particularly SOC and available phosphorus, which are key for soil fertility.

The integration of hyperspectral data and machine learning models XGBoost, Random Forest, and LightGBM proved effective in predicting soil colour and properties. These models offer promising tools for precision agriculture and soil management.

GIS-based soil colour maps generated from the data can provide valuable insights into soil characteristics, aiding in soil mapping. Overall, the study underscores the potential of using soil colour parameters as a simple and effective indicator of soil health.

**Acknowledgments:**

This research was supported by the National Agriculture Higher Education Project, Centre for Advanced Agriculture Science and Technology (NAHEP-CAAST) on "Skill Development to Use Data for Natural Resources Management in Agriculture", College of Agricultural Engineering Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur for providing financial, and the Department of Soil science and Agricultural Chemistry, laboratory facility and other support.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

## Reference


1. Al-Naji, A., Fakhri, A. B., Gharghan, S. K., and Chahl, J. (2021). Soil color analysis based on a RGB camera and an artificial neural network towards smart irrigation: A pilot study. *Heliyon*, 7(1).
2. Arya, S., Singh, R., Jat, M. K., and Yadav, P. K. (2024). Characterization and Classification of Ber (*Ziziphusmauritiana*) Growing Soils of Rewari District, Haryana: Characterization and classification of ber growing soils. Journal of Soil Salinity and Water Quality, 16(1), 1-15.
3. Brady, N. C., and Weil, R. R. (2008). *The nature and properties of soils* (14th ed.). Pearson Education.
4. Conklin, A. R. (2013). *Introduction to soil chemistry: Analysis and instrumentation*. John Wiley and Sons.
5. de Souza, W. M., Ribeiro, A. J. A., and da Silva, C. A. U. (2022). Use of ANN and visual-manual classification for prediction of soil properties for paving purposes. International Journal of Pavement Engineering, 23(5), 1482-1490.
6. Dinesh, Bhat, M.A. and Grewal K.S. (2017) Characterization and classification of soils on different geomorphic units of North–Eastern Haryana, India. Agropedology, 27(02): 103–116.
7. Djama, Z. A., Kavaklıgil, S. S., andErşahin, S. (2023). Evaluation of Soil Color and Soil Fertility Relations on Cultivated Semi-Arid Sloping Landscapes. Journal of Agricultural Faculty of Gaziosmanpaşa University (JAFAG), 40(1), 19-25.
8. Domínguez Soto, J. M., Román Gutiérrez, A. D., Prieto García, F., and Acevedo Sandoval, O. (2012). Munsell notation system and cielab as a tool for evaluation colors in soils. *Revistamexicana de cienciasagrícolas*, *3*(1), 141-155.
9. Ibáñez-Asensio, S., Marques-Mateu, A., Moreno-Ramón, H., and Balasch, S. (2013). Statistical relationships between soil colour and soil attributes in semiarid areas. *Biosystems Engineering*, *116*(2), 120-129.
10. Kang, Y. G., Lee, J. Y., Kim, J. H., and Oh, T. K. (2024). Quantifying soil organic matter for sustainable agricultural land management with soil color and machine learning technique. Agronomy Journal.
11. Kirillova, N. P., Grauer-Gray, J., Hartemink, A. E., Sileova, T. M., Artemyeva, Z. S., and Burova, E. K. (2018). New perspectives to use Munsell color charts with electronic devices. Computers and Electronics in Agriculture, 155, 378-385.
12. Lin, X., Su, Y. C., Shang, J., Sha, J., Li, X., Sun, Y. Y., and Jin, B. (2019). Geographically weighted regression effects on soil zinc content hyperspectral modeling by applying the fractional-order differential. Remote Sensing, 11(6), 636.
13. Liu, F., Rossiter, D.G., Zhang, G. L. and Li, D.C., (2020). A soil colour map of China. Geoderma 379, 114556.
14. Minh, V. Q., Vu, P. T., Du, T. T., and Tinh, T. K. (2023). The Soil Color as Indicator for Fruit Garden Soil Assessment and Recommendation. Agricultural Engineering International: CIGR Journal, 25(2).
15. Moritsuka, N., Matsuoka, K., Katsura, K., and Yanai, J. (2019). Farm-scale variations in soil color as influenced by organic matter and iron oxides in Japanese paddy fields. *Soil Science and Plant Nutrition*, *65*(2), 166-175.
16. Moritsuka, N., Matsuoka, K., Katsura, K., Sano, S., andYanai, J. (2014). Soil color analysis for statistically estimating total carbon, total nitrogen and active iron contents in Japanese agricultural soils. Soil science and plant nutrition, *60*(4), 475-485.

17. N.P. Kirillova (2018) new perspectives to use Munsell color charts with electronic devices Comput. Electron. Agric.
18. Nair, P. R., Nair, V. D., Kumar, B. M., and Haile, S. G. (2009). Soil carbon sequestration in tropical agroforestry systems: a feasibility appraisal. *Environmental Science and Policy*, *12*(8), 1099-1111.
19. Pachlaniya, N. K., Jain, R. K., and Shukla, A. K. (2023). Evaluations of soil fertility status of available major nutrients and micro nutrients in vertisol of indore district of Madhya Pradesh India. Plant archives, 23(1), 100-104.
20. Pegalajar, M. C., Ruíz, L. G. B., Sánchez-Maranon, M., andMansilla, L. (2020). A Munsellcolour-based approach for soil classification using Fuzzy Logic and Artificial Neural Networks. *FUZZY sets and systems*, *401*, 38-54.
21. Poppiel, R.R., Lacerda, M.P.C., Rizzo, R., Safanelli, J.L., Bonfatti, B.R., Silvero, N.E.Q. and Dematte, J.A.M., (2020). Soil color and mineralogy mapping using proximal and remote sensing in Midwest Brazil. Remote Sens. 12, 1197.
22. Ramos, P. V., Inda, A. V., Barrón, V., Siqueira, D. S., Júnior, J. M., and Teixeira, D. D. B. (2020). Color in subtropical brazilian soils as determined with a Munsell chart and by diffuse reflectance spectroscopy. *Catena*, *193*, 104609.
23. Rizzo, R., Wadoux, A. M. C., Demattê, J. A., Minasny, B., Barrón, V., Ben-Dor, E., Francos, N., Savin, I., poppiel, R., silvero,N. E. Q., Silva terra, F. D., Rosin, N. A., Rosas, J. T. F., Greschuk, L. T., Ballester, M. V. R., Das, B. S., Malone, B. P. and Salama, E. S. M. (2023). Remote sensing of the Earth's soil color in space and time. *Remote Sensing of Environment*, *299*, 113845.
24. Sawut, R., Kasim, N., Abliz, A., Hu, L., Yalkun, A., Maihemuti, B., andQingdong, S. (2018). Possibility of optimized indices for the assessment of heavy metal contents in soil around an open pit coal mine area. International journal of applied earth observation and geoinformation, 73, 14-25.
25. Schmidt, M. W., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., and Trumbore, S. E. (2011). Persistence of soil organic matter as an ecosystem property. *Nature*, *478*(7367), 49-56.
26. Schmidt, S.A., Ahn, C., 2021. Analysis of soil color variables and their relationships between two field-based methods and its potential application for wetland soils. Sci. Total Environ. 783, 147005
27. Schwertmann, U., and Taylor, R. M. (1989). Iron oxides. In Soil chemistry. Springer. pp. 145-170.
28. Silvero, N. E. Q., Demattê, J. A. M., Amorim, M. T. A., dos Santos, N. V., Rizzo, R., Safanelli, J. L. and Bonfatti, B. R. (2021). Soil variability and quantification based on Sentinel-2 and Landsat-8 bare soil images: A comparison. Remote Sensing of Environment, 252, 112117.
29. ViscarraRossel, R.A., Bui, E.N., De Caritat, P. and McKenzie, N.J., (2010). Mapping iron oxides and the color of australian soil using visible-near-infrared reflectance spectra. J. Geophys. Res. Earth Surf. 115, 1–13.
30. Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., ... and Kögel-Knabner, I. (2019). Soil organic carbon storage as a key function of soils-A review of drivers and indicators at various scales. *Geoderma*, *333*, 149-162.