Int. Agrophysics, 2004, 18, 1-10

INTERNATIONAL
Agrophysics
www.ipan.lublin.pl/int-agrophysics

Soil colour and spectral analysis employing linear regression models
I. Effect of organic matter

P.E. Barouchas and N.K. Moustakas*

Agricultural University of Athens, Laboratory of Soil Science and Agricultural Chemistry, 75 Iera Odos, Botanikos 11855,
Athens, Greece

Received April 1, 2003; accepted October 6, 2003

Abstract. This work comprises an investigation into whether soil reflectance spectral analysis which is employed to calculate the colour characteristics (hue, value, chroma) of soil can be carried out using linear regression models, so that comparison of colour characteristics subsequently becomes possible, and also statistically documented. To this end the colour of soil samples was calculated through spectrum reflectance in the visible region of dry smooth-rubbed soil samples smaller than 250 μm. The colour parameters of the CIE system assessed by analysis of the spectrum reflectance were converted into Munsell colour system characteristics. Regression in accordance with the piecewise linear model was then applied to the spectrum data. The processing indicated that this model is capable of making satisfactory predictions – above all of the value and secondarily of the chroma of the soil samples. Detection of statistically significant differences in the colour characteristics of horizons of the same profile was effected through the application of the nested model. These differences cannot be detected using the tables of the Munsell colour system. Finally, in each region of the spectrum, qualitative analysis of the effect of the organic matter on the soil colour characteristics was performed, demonstrating its active role in determining the readings for value and chroma.

Keywords: soil colour, spectral analysis, organic matter

INTRODUCTION

Collecting soil reflectance spectra and subjecting them to analysis is a contemporary technique of soil science for the study of soil properties. Spectrum data have been used to calculate the colour of soil [5,10,20], the effect of iron oxides on soil colour [3,8] and the effect of organic matter on the reflectance properties of soils [6,17]. Moreover, changes in the colour characteristics of soils have been used for the study of their genesis [1,10,14,20]. Organic matter is an important component of soil, and numerous researchers [4,13] have noted its role in the determination of colour. To this day differences in colour characteristics of soil profiles are described through the use of the tables of the Munsell colour system. Because the Munsell colour system is not able to accurately convey these differences, the use of spectroscopic methods for calculating the colour characteristics is an indispensable tool for scientists seeking precise characterization of colour.

The purpose of this study is to examine whether the colour characteristics which could be estimated from reflectance data of soil samples may be expressed in a linear fashion, rendering it possible subsequently for them to be subjected to statistically documented comparison. There is also an investigation, using the models developed, concerned with which of the colour characteristics the organic matter reacts to, and to what extent.

MATERIALS AND METHODS

Sampling

The soil samples are taken from Alfisols-order soils [15] from the region of Thessaly (Central Greece) and are representative alluvial deposits of the quaternary period. The soil samples OL1 to OL8 and D1 to D15 were taken at every 15 cm, while the samples EL1 to EL6 and PL1 to PL3 were taken in accordance with the depth of the soil horizons. The soil samples were air-dried, smooth-rubbed and sifted with a 250 μm sieve [5]. To ascertain the role of organic matter in determining the colour characteristics (hue, value and chroma), soil spectrum reflectance was carried out in the visible region (380-770 nm), both on untreated soil samples

---

*Corresponding author's e-mail: barouchas@agr.forthnet.gr and after the removal of the organic matter. This was effected by means of 30% H_2O_2 [12]. The soil spectrum reflectance was carried out in accordance with the procedure recommended by Fernandez [5] and Torrent [18]. Calculation of soil colour then took place on the basis of the CIE system [7], converting them into colour coordinates x, y, z [2,19]. The reflectance spectra of the soil samples were recorded with a Perkin-Elmer Lambda 15 UV/VIS spectrophotometer equipped with an integrated sphere. The spectra were measured from 380 to 770 nm in 5 nm step. The colour conversion from the CIE system to the Munsell colour system was carried out with a GretagMacbeth Company software program. The hue values were set out on a scale from 0.05 for a hue reading of 0.01YR to 100 for a hue reading of 10Y with a step of 0.05.

Linear regression statistical models

Piecewise linear regression

Regression of Y data with X data often proceeds in accordance with a special linear relationship in a certain X field, but can show a different linear relationship in some other fields of X [11]. In the case of soil reflectance spectrum, the Y values represent the reflectance radiation expressed on a scale from 0 for zero reflection or complete absorption to 1 for total reflection or zero absorption. The X values represent the wavelength of visible region (380-770) expressed in nm. The region of the spectrum from 380 to 770 nm can be divided into three regions, a) 380-500 nm (the blue region of the spectrum), b) 501-600 nm (the green region of the spectrum), c) 601-770 nm (the red region of the spectrum). In each of these areas, the spectrum is governed by a different linear relationship which can be described by a function, the general formula for which is:

$$Y = b_o+b_1X+b_2\,(X\text{-}500)\,X_1+b_3\,(X–600)\,X_2, \quad (1)$$

where: X is the wavelength of the visible radiation (380–770 nm), X_1 is 1 when X>500 or X_1 is 0 when X< 500, X_2 is 1 when X>600 or X_2 is 0 when X< 600, b_1, b_2, b_3 are constants. Equation (1) takes X values from 380 to 500 nm, then $X_1 = X_2 = 0$, and accordingly (1) becomes:

$$Y = b_o+b_1\,X, \quad (2)$$

where: b_0 denotes the constant term (point of intersection of the line with the Y axis) and b_1 denotes the slope of the line. Wherever Eq. (1) takes X values from 501 to 600 nm, then $X_1 = 1$, $X_2 = 0$ and accordingly (1) becomes:

$$Y = b_o+b_1X+b_2\,(X–500)\,X_1 = (b_0–500\,b_2)+(b_1+b_2)\,X. \quad (3)$$

The formula (b_0–500 b_2) denotes the constant term (point of intersection of the line with the Y axis) and the formula (b_1+b_2) denotes the slope of the line. Wherever Eq. (1) takes X values from 601 to 770 nm, then $X_1 = 0$, $X_2 = 1$ and accordingly (1) becomes:

$$Y = b_o+b_1\,X+b_3\,(X–600)\,X_2 = (b_0–600\,b_3)+(b_1+b_3)\,X. \quad (4)$$

The formula (b_0–600 b_3) denotes the constant term (point of intersection of the line with the Y axis) and the formula (b_1+b_3) denotes the slope of the line.

Multiple linear regression

The general description of a linear model with two or more variables can be conveyed by the Eq. (5):

$$Y_i = b_0+b_1X_{i1}+b_2\,X_{i2}+\ldots.+b_{p–1}\,X_{i,p–1}+{}_i\,, \quad (5)$$

where: p–1 is the number of the variables X, b_o, b_1,…$b_{p–1}$ are parameters, and ${}_i$ is the random error. By definition the mean value for random error E{ ${}_i$} is equal to 0, and consequently the function which describes the change in the mean value Y_i(E{Y_i}), is:

$$E\{Y_i\} = b_0+b_1\,X_{i1}+b_2\,X_{i2}+\ldots.+b_{p–1}\,X_{i,p–1}. \quad (6)$$

Function (6) has more than two dimensions. The significance, for example, of the cofficient X_{i1}, i.e., b_1, is that it denotes the change in the mean value E{Y_i} for one unit of change in the value X_{i1}, keeping all the other variables constant.

Nested Model

Two mathematical models are called nested when one contains all the terms of the other and at least one term more [9]. Let us postulate that someone wishes to correlate the reflected radiation (Q) of two or more soil samples (soil samples constitute independent units) coming from one or more soil profiles of the same soil order, with the corresponding wavelength of radiation. In this instance the questions which arise are to what extent the Q of the two samples is different or whether the rate of increase of Q is different for two or more soil samples. The answer to these questions can be provided through the use of a nested linear regression model characterizing Q as a function of two independent variables: one quantitative – the wavelength, and one qualitative – the soil sample. The development of a nested linear regression model is described as follows:

The linear relationship between the proportional reflectance Q (expressed on a scale from 0 to 1) and the wavelength of the reflected radiation is the same for all the soil samples:

$$Q = b_0+b_1\,X_1+b_2\,X_2+b_3\,X_3, \quad \text{(Model 1)}$$

where $380<X_1<500$, $501<X_2<600$ $601<X_3<770$ nm. Model 1 represents the mean value for the soil samples under examination.

The linear relationship between the proportional reflectance Q (expressed on a scale from 0 to 1) and the wavelength of the reflected radiation is different for each soil sample, while the rate of increase of Q for each 1 nm change in the wavelength is the same for all the soil samples under examination:

$$Q= b_0+b_1X_1+b_2X_2+b_3X_3+\ldots+b_{3n-2}\, X_{3n-2}+ b_{3n-1}X_{3n-1}+b_{3n}\, X_{3n}, \qquad \text{(Model 2)}$$

where: X_1, X_2, X_3, are the wavelengths and X_{3n-2}, X_{3n-1}, X_{3n} are the dummy variables which correspond to the values: $X_4 = X_5 = X_6 = 1$, when sample 2 is examined and $X_4 = X_5 = X_6 = 0$ when it is not examined, $X_7 = X_8 = X_9 = 1$, when sample 3 is examined and $X_7 = X_8 = X_9 = 0$ when it is not examined. $X_{3n-2} = X_{3n-1} = X_{3n} = 1$, when the sample n is examined and $X_{3n-2} = X_{3n-1} = X_{3n} = 0$ when it is not examined, and n = 2, 3,…. This model is a combination of two other separate models, a first-order linear model ($Q = b_0+b_1X_1+b_2\, X_2+b_3\, X_3$) with three quantitative variables (X_1, X_2, X_3) and a first-order model ($Q=b_0+b_1\, X_1+b_2\, X_2+b_3\, X_3+\ldots+b_{3n-2}\, X_{3n-2}+b_{3n-1}\, X_{3n-1}+b_{3n}X_{3n}$) with three or more qualitative variables (X_{3n-2}, X_{3n-1}, X_{3n}).

The second model presupposes that there is no interaction between the quantitative and qualitative variables. To describe to what extent the linear equations correlating the reflectance and the wavelength differ over two or more soil samples, i.e., how far the values for the points of intersection with the Y axis (intercepts) and the slopes in one of the linear equations differ from those in the other, the following model is used:

$$Q = b_0+b_1\, X_1 + b_2\, X_2 + b_3\, X_3 + b_{3n-2}\, X_{3n-2} + b_{3n-1}\, X_{3n} + b_{3n}\, X_{3n} + b_{3n+1}\, X_1\, X_4 + b_{3n+2}\, X_2\, X_5 + b_{3n+3}\, X_3\, X_6 + b_{3n+4}\, X_1\, X_7 + b_{3n+5}\, X_2\, X_8 + b_{3n+6}\, X_3\, X_9 +\ldots+ b_{k-2}\, X_1\, X_{3n-2} + b_{k-1}\, X_2\, X_{3n-1} + b_k\, X_3\, X_{3n}, \qquad \text{(Model 3)}$$

where: k is the number of b parameters other than b_0 and n = 2,3,…. The terms denoting the quantitative variables X_1, X_2, X_3 and the qualitative variables X_{3n-2}, X_{3n-1}, X_{3n} correspondingly are called main effect terms, while the terms $X_1\, X_{3n-2}$, $X_2\, X_{3n-1}$, $X_3\, X_{3n}$ are called 'interaction terms'.

Each of the three models described was created by adding terms to Model 1. Model 2 was created by adding the main effect terms to Model 1; Model 3 was created by adding the interaction terms to Model 2. These models are nested (Model 3 is nested in Models 1 and 2; Model 2 is nested in Model 1) and can, therefore, be compared using the F-test for nested models [9], since our observations follow a regular distribution. In the event that we wish to compare two nested models, for example Model 1 and Model 3, the question is whether at least one of the parameters b_{3n+1}, b_{3n+2}, b_{3n+3},…, b_k differs from 0. The comparison in this case is made using the F_{calc}, which is calculated according to the formula:

$$F_{calc} = ((SSE_1 - SSE_3)/(k-1))/(SSE_3/(n-(k+1))) = ((SSE_1 - SSE_3)/k-1)/MSE_3,$$

where: SSE_1 is the sum of squared residuals for the Model 1, SSE_3 is the sum of squared residuals for the Model 3, k–3 is the number of b parameters mentioned in the null hypothesis: Ho: $b_{3n+1} = b_{3n+2} = b_{3n+3}= \ldots= b_k = 0$, k+1 is the number of b parameters in Model 3 including the parameter b_0, n is the total sample size

The rejection region for the null hypothesis is for $F_{calc}>F_{a(table)}$, where the value $F_{a(table)}$ is found from tables for v_1 = k–3 degrees of freedom of the numerator and v_2 = n–(k+1) degrees of freedom of the denominator, at level of significance a.

The statistical processing was carried out using the statistical software STATISTICA^{TM} [16].

RESULTS AND DISCUSSION

Application of the piecewise linear regression model

Initially the piecewise linear regression model was applied to the spectrum data with a view to examining whether the sigmoid spectrum reflectance curve could be plotted in a linear fashion. The results, both prior to and subsequent to the removal of the organic matter, for all three regions (380-500 nm, 501-600 nm and 601-770 nm) of the visible spectrum, are indicated on Tables 1 and 2. In all the samples the linear regression coefficient (R^2) between the reflectance data and the wavelength is equal to 0.99.

The reflectance spectra emerging from the application of the piecewise linear regression were used for re-specification of the colour characteristics (predicted values). Conversion of the data into the Munsell colour system characteristics took place in accordance with the E309-96 procedure [2]. The results appear in Tables 3 and 4. We note that there is a constant underestimation of the hue, both before and after the removal of the organic matter. The possibility of predicting the colour value, using the piecewise linear regression model, both before and after the removal of the organic matter, is very high (R^2=0.98). However, the possibility of predicting the values for chroma employing the same model is smaller, both before (R^2=0.885) and after the removal of the organic matter (R^2=0.874), but nevertheless it remains high.

Application of the nested linear regression model

Subsequently, an attempt was made to compare the reflectance spectra which were obtained from the application of the piecewise linear regression model. The comparison was carried out through application of the nested linear regression model. The results appear in Table 5. The SSE1

Table 1. Slopes of predicted reflectance data in the three visible spectral regions and linear regression coefficient (R^2) between reflectance and wavelength, in soil samples without treatment after the implementation of piecewise linear regression model ($Y= b_o+b_1 X+b_2 (X-500) X_1+b_3 (X-600) X_2$)

Sample	380-400 nm		401-500 nm		501-770 nm	
	Slope	R^2	Slope	R^2	Slope	R^2
D1	0.011	0.99	0.130	0.99	0.080	0.99
D2	0.010	0.99	0.148	0.99	0.090	0.99
D3	0.011	0.99	0.149	0.99	0.092	0.99
D4	0.008	0.99	0.156	0.99	0.096	0.99
D5	0.008	0.99	0.166	0.99	0.100	0.99
D6	0.010	0.99	0.150	0.99	0.093	0.99
D7	0.006	0.99	0.109	0.99	0.078	0.99
D8	0.005	0.99	0.121	0.99	0.085	0.99
D9	0.007	0.99	0.125	0.99	0.087	0.99
D10	0.006	0.99	0.142	0.99	0.094	0.99
D11	0.006	0.99	0.147	0.99	0.095	0.99
D12	0.013	0.99	0.182	0.99	0.104	0.99
D13	0.021	0.99	0.160	0.99	0.089	0.99
D14	0.029	0.99	0.191	0.99	0.095	0.99
D15	0.032	0.99	0.182	0.99	0.091	0.99
EL1	0.060	0.99	0.141	0.99	0.070	0.99
EL2	0.053	0.99	0.186	0.99	0.077	0.99
EL3	0.060	0.99	0.191	0.99	0.072	0.99
EL4	0.059	0.99	0.169	0.99	0.072	0.99
EL5	0.072	0.99	0.179	0.99	0.071	0.99
EL6	0.101	0.99	0.166	0.99	0.059	0.99
PL1	0.063	0.99	0.110	0.99	0,065	0.99
PL2	0.034	0.99	0.111	0.99	0.073	0.99
PL3	0.043	0.99	0.117	0.99	0.067	0.99
OL1	0.017	0.99	0.072	0.99	0.056	0.99
OL2	0.015	0.99	0.083	0.99	0.059	0.99
OL3	0.018	0.99	0.113	0.99	0.070	0.99
OL4	0.016	0.99	0.107	0.99	0.068	0.99
OL5	0.014	0.99	0.108	0.99	0.067	0.99
OL6	0.014	0.99	0.115	0.99	0.070	0.99
OL7	0.015	0.99	0.118	0.99	0.070	0.99
OL8	0.022	0.99	0.129	0.99	0.075	0.99

Table 2. Slopes of predicted reflectance data in the three visible spectral regions and coefficient of linear regression coefficient (R^2) between reflectance and wavelength, in soil samples after the removal of the organic matter and the implementation of piecewise linear regression model ($Y= b_o+b_1 X+b_2 (X-500) X_1+b_3 (X-600) X_2$)

Sample	380-400 nm		401-500 nm		501-770 nm	
	Slope	R^2	Slope	R^2	Slope	R^2
D1	0.012	0.99	0.179	0.99	0.093	0.99
D2	0.012	0.99	0.166	0.99	0.092	0.99
D3	0.009	0.99	0.161	0.99	0.094	0.99
D4	0.007	0.99	0.151	0.99	0.095	0.99
D5	0.011	0.99	0.169	0.99	0.097	0.99
D6	0.012	0.99	0.168	0.99	0.097	0.99
D7	0.007	0.99	0.155	0.99	0.096	0.99
D8	0.009	0.99	0.163	0.99	0.096	0.99
D9	0.013	0.99	0.197	0.99	0.099	0.99
D10	0.011	0.99	0.186	0.99	0.100	0.99
D11	0.011	0.99	0.193	0.99	0.107	0.99
D12	0.015	0.99	0.204	0.99	0.105	0.99
D13	0.026	0.99	0.215	0.99	0.101	0.99
D14	0.036	0.99	0.223	0.99	0.094	0.99
D15	0.037	0.99	0.217	0.99	0.091	0.99
EL1	0.072	0.99	0.181	0.99	0.066	0.99
EL2	0.075	0.99	0.214	0.99	0.065	0.99
EL3	0.097	0.99	0.244	0.99	0.055	0.99
EL4	0.077	0.99	0.246	0.99	0.068	0.99
EL5	0.094	0.99	0.204	0.99	0.064	0.99
EL6	0.075	0.99	0.219	0.99	0.075	0.99
PL1	0.095	0.99	0.146	0.99	0.050	0.99
PL2	0.057	0.99	0.155	0.99	0.066	0.99
PL3	0.101	0.99	0.161	0.99	0.071	0.99
OL1	0.028	0.99	0.194	0.99	0.081	0.99
OL2	0.021	0.99	0.191	0.99	0.088	0.99
OL3	0.016	0.99	0.183	0.99	0.089	0.99
OL4	0.020	0.99	0.168	0.99	0.076	0.99
OL5	0.022	0.99	0.163	0.99	0.072	0.99
OL6	0.020	0.99	0.150	0.99	0.073	0.99
OL7	0.018	0.99	0.141	0.99	0.071	0.99
OL8	0.021	0.99	0.146	0.99	0.079	0.99

Table 3. Observed and predicted soil colour attributes using the piecewise linear regression model for soil samples without treatment

Sample	Observed values[a]				Predicted values[b]			
	Hue	Scale[c]	Value	Chroma	Hue	Scale[c]	Value	Chroma
D1	4.73YR	23.65	4.38	3.81	3.64YR	18.20	4.32	3.72
D2	4.49YR	22.45	4.52	4.17	3.56YR	17.80	4.39	4.04
D3	4.57YR	22.85	4.50	4.25	3.13YR	15.65	4.42	3.86
D4	4.22YR	21.10	4.58	4.29	3.16YR	15.80	4.39	4.06
D5	4.15YR	20.75	4.72	4.41	3.20YR	16.00	4.47	4.32
D6	4.44YR	22.20	4.54	4.21	3.18YR	15.90	4.45	4.26
D7	4.42YR	22.10	4.18	3.34	2.92YR	14.60	4.12	3.25
D8	4.19YR	20.95	4.33	3.55	2.97YR	14.85	4.19	3.19
D9	4.43YR	22.15	4.27	3.76	3.05YR	15.25	4.32	3.71
D10	4.09YR	20.45	4.46	4.03	2.91YR	14.55	4.30	3.68
D11	4.12YR	20.60	4.47	4.16	2.97YR	14.85	4.36	3.80
D12	4.45YR	22.25	4.80	4.78	3.19YR	15.95	4.77	4.32
D13	5.17YR	25.85	4.75	4.33	4.16YR	20.80	4.70	4.23
D14	5.26YR	26.30	5.12	4.73	4.38YR	21.90	5.10	4.34
D15	5.49YR	27.45	5.11	4.55	4.72YR	23.60	4.98	4.37
EL1	7.81YR	39.05	5.55	3.38	6.34YR	31.70	5.45	2.9
EL2	6.58YR	32.90	5.50	4.32	4.64YR	23.20	5.37	3.81
EL3	6.91YR	34.55	5.62	4.34	4.84YR	24.20	5.48	4.04
EL4	7.11YR	35.55	5.79	3.67	5.06YR	25.30	5.46	3.30
EL5	7.56YR	37.80	5.99	3.79	5.98YR	29.90	5.90	3.68
EL6	8.78YR	43.90	6.31	3.50	7.82YR	39.10	6.26	3.34
PL1	8.64YR	43.20	5.65	2.68	8.19YR	40.95	5.74	2.34
PL2	6.89YR	34.45	4.83	3.11	5.43YR	27.15	4.97	2.53
PL3	7.51YR	37.55	5.07	3.13	6.37YR	31.85	5.12	2.72
OL1	6.42YR	32.10	4.05	2.41	6.47YR	32.35	4.18	2.55
OL2	5.97YR	29.85	4.08	2.71	5.64YR	28.20	4.40	2.54
OL3	5.52YR	27.60	4.37	3.41	5.44YR	27.20	4.43	3.30
OL4	5.51YR	27.55	4.25	3.32	5.35YR	26.75	4.36	3.28
OL5	5.31YR	26.55	4.29	3.26	5.33YR	26.65	4.42	3.43
OL6	5.19YR	25.95	4.36	3.41	5.15YR	25.75	4.47	3.57
OL7	5.18YR	25.90	4.41	3.44	5.14YR	25.70	4.50	3.24
OL8	5.45YR	27.55	4.74	3.51	5.28YR	26.40	4.92	3.68

[a] Soil colour characteristics estimated from the original spectral data, [b] Soil colour characteristics estimated after the use of piecewise linear regression model, [c] Hue scale from 0.05 for Hue 0.01YR to 100 for Hue 10Y with 0.05 step.

Table 4. Observed and predicted soil colour attributes using the piecewise linear regression model for soil samples after soil organic matter destruction

Sample	Observed values[a]				Predicted values[b]			
	Hue	Scale[c]	Value	Chroma	Hue	Scale[c]	Value	Chroma
D1	4.38YR	21.90	4.85	4.60	3.07YR	15.35	4.76	4.36
D2	4.53YR	22.65	4.65	4.51	3.17YR	15.85	4.57	4.28
D3	4.33YR	21.65	4.63	4.37	2.98YR	14.90	4.55	4.17
D4	4.20YR	21.01	4.51	4.25	2.76YR	13.80	4.43	4.02
D5	4.46YR	22.30	4.68	4.57	3.10YR	15.50	4.59	4.37
D6	4.43YR	22.15	4.75	4.45	3.08YR	15.40	4.68	4.25
D7	4.25YR	21.25	4.53	4.32	2.86YR	14.30	4.45	4.14
D8	4.29YR	21.45	4.69	4.36	2.96YR	14.80	4.61	4.16
D9	4.29YR	21.45	5.03	4.81	3.03YR	15.15	4.94	4.62
D10	4.34YR	21.70	4.84	4.80	3.02YR	15.10	4.75	4.58
D11	4.35YR	21.75	4.77	5.09	2.94YR	14.70	4.67	4.83
D12	4.52YR	22.60	4.94	5.18	3.21YR	16.05	4.86	4.92
D13	5.02YR	25.10	5.13	5.27	3.92YR	19.60	5.05	5.08
D14	5.39YR	26.95	5.40	5.13	4.45YR	22.25	5.33	4.95
D15	5.34YR	26.70	5.54	4.78	4.48YR	22.40	5.48	4.62
EL1	7.53YR	37.65	6.07	3.73	6.50YR	32.50	5.77	3.92
EL2	7.12YR	35.60	6.13	4.30	5.74YR	28.70	5.87	4.28
EL3	7.57YR	37.85	6.43	4.67	6.18YR	30.90	6.15	5.02
EL4	6.71YR	33.55	6.33	4.66	5.21YR	26.05	6.01	4.87
EL5	7.87YR	39.35	6.47	3.93	7.03YR	35.15	6.20	4.28
EL6	6.96YR	34.80	6.24	4.29	5.74YR	28.70	5.91	4.38
PL1	9.04YR	45.20	6.43	2.97	8.75YR	43.75	6.43	2.93
PL2	7.49YR	37.45	5.41	3.78	6.72YR	33.60	5.39	3.73
PL3	5.78Y	28.90	6.40	3.41	6.97YR	34.85	5.31	3.62
OL1	5.27YR	26.35	5.11	4.77	4.06YR	20.30	5.05	4.59
OL2	4.88YR	24.40	4.96	4.84	3.62YR	18.10	4.88	4.65
OL3	4.68YR	23.40	4.86	4.73	3.36YR	16.80	4.77	4.53
OL4	4.68YR	23.40	4.86	4.73	3.88YR	19.40	4.70	4.29
OL5	5.26YR	26.30	4.81	4.30	4.05YR	20.25	4.74	4.16
OL6	5.29YR	26.45	4.64	4.16	4.10YR	20.50	4.58	4.03
OL7	5.22YR	26.10	4.57	3.97	3.96YR	19.80	4.51	3.85
OL8	5.01YR	25.05	4.88	3.79	3.90YR	19.50	4.83	3.69

Explanations as in Table 3.

and SSE3 values represent the sum of squared residuals for Eqs (1) and (3), respectively. The SSE1-SSE3 difference can be regarded as the measure of the difference between the reflectance spectra described through the nested linear regression model [9]. The absolute difference in predicted colour value as between soil horizons is symbolized by the d-Value. The critical $F_{0.05}$ reading is 1.88. In all these comparisons the F value is higher than the critical value, which means that there is a statistically significant difference between the reflectance spectra of the horizons under examination. In most cases the d-Value is less than a unit (Table 5). A difference of this kind cannot be registered through colour measurement either directly from Munsell colour tables or through conversion of spectrum data into Munsell colour characteristics. The high value for the linear regression coefficient (R^2= 0.923), between the SSE1-SSE3 difference and the d-Value shows that the nested linear regression model successfully registers differences in colour value between soil horizons of the same soil order.

Application of the models developed.

Effect of organic matter

In this study an attempt is made, employing the linear models that have been elaborated, to determine what colour characteristics are affected by organic matter, and to what extent. With a view to answering the first part of the question, simple correlations were drawn between the observed colour characteristics and the predicted slope of the reflectance spectra emerging from the employment of the piecewise linear regression model in the three regions of the visible spectrum (380-500, 501-600 and 601-770 nm). The results are shown in Table 6. On the assumption that change in the slope of the predicted reflectance spectrum is affected in a linear fashion by the organic matter, with a view to answering the second part of the question, multiple linear regression (stepwise) was employed between the slope of the predicted reflectance spectrum as an independent variable and the observed colour parameters as independent

Table 5. F values of predicted colour characteristic value between soil horizons, for the P1, P2, P3 and P4 profiles

Profile	Horizon	Treatment	SSE1[a]	SSE3[b]	F[c]	SSE1-SSE3[d]	d-Value[e]
P_1	D1-D7	NT[f]	0.1287	0.0441	247.5	0.085	0.20
		OMD[g]	0.2651	0.0988	216.5	0.166	0.31
P_2	E1-E3	NT[f]	0.1555	0.0109	1721.3	0.145	0.03
		OMD[g]	0.6063	0.0214	3481.7	0.360	0.38
P_3	P1-P3	NT[f]	0.5897	0.0037	19533.4	0.586	0.62
		OMD[g]	0.8649	0.0090	11887.0	0.856	1.12
P_4	OL1-OL4	NT[f]	0.1999	0.0158	1534.4	0.184	0.18
		OMD[g]	0.3662	0.0832	436.6	0.283	0.35

[a] Sum of the squared residuals for the Model 1, [b] Sum of the squared residuals for the Model 3, [c] F value for comparing spectral data, [d] SSE3-SSE1difference is the weight of the difference between spectral data, [e] Absolute difference of value characteristic between soil horizons, [f] Sample without treatment, [g] Sample after soil organic matter destruction.

Table 6. Simple correlation coefficient (r) between the slopes of predicted reflectance data and the observed soil colour characteristics (hue, value and chroma), before and after the removal of the soil organic matter in three visible regions

Spectral band (nm)	Before removal of organic matter			After the removal of organic matter		
	Hue (Scaled)	Value	Chroma	Hue (Scaled)	Value	Chroma
380-500	0.7437[a]	0.9688[a]	-0.2524	0.9407[a]	0.9245[a]	-0.2186
501-600	-0.2518	0.4808[a]	0.7575[a]	-0.0633	0.3674[a]	0.7397[a]
601-770	-0.8437[a]	-0.3080	0.7853[a]	-0.8577[a]	-0.7083[a]	0.4700[a]

[a] Significant at P = 0.05.

Table 7. Multiple linear regress ion functions between the slopes of predicted reflectance data and the observed colour characteristics (hue, value, chroma), prior to and after the removal of the soil organic matter in three visible spectral regions

Visible region (nm)	Before the removal of organic matter	R^2	After the removal of organic matter	R^2
	Regression equations		Regression equations	
380-500	$S_{(380\text{-}500)}{}^{b}=0.068-0.0012* H_o{}^{e}+0.011* C_o{}^{g}$	0.84^a	$S_{(380\text{-}500)}= 0.1209-0.0017* H_o$	0.95^a
501-600	$S_{(501\text{-}601)}{}^{c}= -0.291-0.0057*H_o{}^{e}+0.093*V^{f}+0.028*C_o{}^{g}$	0.95^a	$S_{(501\text{-}601)}=-0.211-0.0016* H_o+0.048*V_o$	0.95^a
601-770	$S_{(601\text{-}770)}{}^{d}=-0.172+0.042* V_o{}^{f}$	0.94^a	$S_{(601\text{-}770)}=-0.0471+0.004* H_o$	0.88^a

[a] Significant at P= 0.05, [b] Slope in the 380-500 nm visible region, [c] Slope in the 501-600 nm visible region, [d] Slope in the 601-770 nm visible region, [e] Observed hue (scaled) colour characteristic, [f] Observed value colour characteristic, [g] Observed chroma colour c.

variables in the corresponding region of the spectrum. The results are shown in Table 7.

In the region 380-500 nm, the slope of the predicted linear reflectance spectrum correlates with the observed values for hue and value, both before and after the removal of the organic matter from the soil samples (Table 6). From the application of multiple regression with a dependent variable on the slope of spectrum, with the observed colour parameters as independent variables, it appears that before the removal of the organic matter the slope of the spectrum can be assessed from the colour value, whereas after the removal of the organic matter from the hue (Table 7).

In the region 501-600 nm, the slope of the predicted linear reflectance spectrum correlates with the colour characteristics of value and chroma, both before and after the removal of the organic matter (Table 6). From the application of multiple regression it appears that the slope of the spectrum can be assessed, both before and after the removal of the organic matter, from all the colour characteristics (Table 7).

In the region 601-770 nm, the predicted linear reflectance slope correlates with the colour characteristics hue and chroma before the removal of the organic matter. However, after the removal of the organic matter, the slope of the reflectance spectrum correlates with all the colour characteristics (Table 6). From the application of multiple regression it appears that the slope of the spectrum can be assessed, before the removal of the organic matter, from hue and chroma, but after the removal of the organic matter only from the hue (Table 7).

CONCLUSIONS

1. Concluding, the procedure followed may constitute a qualitative method in examination of the effect of organic matter on the spectrum data and indirectly of the effect of soil colour. The organic matter of the soil contributes to determining the values for value and chroma of soil and to a lesser extent also contributes to determining hue.

2. Application of the piecewise linear regression model makes possible an approach to the visible spectrum of reflected radiation (380-770 nm) in soil samples without this influencing calculation of the value of soil colour, irrespective of the presence or absence of organic matter.

3. Application of the nested linear regression model can provide statistical documentation for differences in readings, primarily for colour value, even smaller than a single unit, between soil samples. The Munsell colour system can not register these differences.

REFERENCES

1. **Allen B.L. and Fanning D.S., 1983.** Composition and soil genesis. In: Pedogenesis and Soil Taxonomy. I. Concepts and Interactions. (Eds L.P. Wilding, N.E. Smeck and G.F. Hall), 141-192. Elsvier, Amsterdam.
2. ASTM, **1996.** Standard Practice for Computing the Colors of Objects by Using the CIE System. ASTM standard E 308-96. American Society for Testing and Materials, Philadelphia.
3. **Barron V. and Torrent J., 1986.** Use of the Kubelka-Munk theory to study the influence of iron oxides on soil colour. J. Soil Sci., 37, 499-510.
4. **Baumgardner M.F., Kristof S.J., Johannsen C.J., and Zachary A.L., 1970.** Effects of organic matter on the multispectral properties of soils. Proc. Indian Academy of Sciences, 79, 413-422.
5. **Fernandez R.N. and Schulze D.G., 1987.** Calculation of soil color from reflectance spectra. Soil Sci. Soc. Am. J., 51, 1277-1282.
6. **Henderson T.L., Baumgardner M.F., Franzmeier D.P., Stott D.E., and Coster D.C., 1992.** High dimensional reflectance analysis of soil organic matter. Soil Sci. Soc. Am. J., 56, 865-872.
7. **Hunt R.W.G., 1998.** Measuring Colour. Fountain Press, England.
8. **Kosmas C.S., Franzmeier D.P. and Schulze D.G., 1986.** Relationship among derivative spectroscopy, color, crystallite dimensions, and Al substitution of synthetic goethites and hematites. Clays and Clay Minerals, 34, 625-634.
9. **Medenhall W. and Sinich T., 1996.** A Second Course in Statistics: Regression Analysis. Prentice Hall, New Jersey.
10. **Moustakas N., 1990.** Relations between morphological and physicochemical properties of Vertisols in Greece. PhD thesis, Agricultural University of Athens, Greece.
11. **Neter J., Kutner M.H., Nachtsheim C.J., and Wasserman W., 1996.** Applied linear statistical models. IRWIN, Inc.

12. **Page A.L., Miller R.H., and Buxton D.R., 1982.** Methods of soil analysis, Part 2. Chemical and microbiological properties. Am. Soc. Agron., Madison.
13. **Schulze D.G., Nagel J.L., Scoyoc G.E., Henderson, T.L., and Baumgardner M.F., 1993.** Significance of organic matter in determining soil colors. In: Soil Colors (Eds J.M. Bigham and E.J. Ciolkosz), 71-90, Soil Sci. Soc. Amer., Madison.
14. **Schwertmann U. and Taylor R.M., 1977.** Iron Oxides. In: Minerals in Soil Environments (Eds J.B. Dixon and S.B. Weed), 145-180, Soil Sci. Soc. Amer., Madison.
15. Soil Survey Staff, Soil Taxonomy, **1975.** A Basic System of Soil Classification for Making and Interpreting Soil Surveys. USDA Handbook No. 436, Soil Conservation Service, U.S. Government Printing Office, Washington, DC.
16. StatSoft Inc., **1995.** Statistica for Windows. Tulsa, OK.
17. **Stoner E.R. and Baumgardner M.F., 1981.** Characteristic variations in reflectance of surface soils. Soil Sci. Soc. Am. J., 45, 1161-1165.
18. **Torrent J. and Barron V., 1993.** Laboratory measurements of soil color: theory and practice. In: Soil Colors (Eds J.M. Bigham and E.J. Ciolkosz), 71-90, Soil Sci. Soc. Am., Madison.
19. **Wyszecki G. and Stiles W.S., 1982.** Color science: Concepts and methods, quantitative data and formulae. Wiley and Sons, New York.
20. **Yassoglou N., Kosmas C. and Moustakas N., 1977.** Red soils: their properties, use and management in Greece. Catena, 28, 261-278.