AJA VOL (3) ISSUE 2, 2016: 45-49

*Azarian Journal of Agriculture*

www.azarianjournals.ir

*Azarian Journals*

***Research article*** **ISSN:2383-4420**

# Soil organic carbon prediction using remotely sensed data at Lighvan watershed, Northwest of Iran

Mehdi Rahmati[1*], Mohammad Reza Neyshabouri[2], Majid Mohammady Oskouei[3], Ahmad Fakheri Fard[4], Abbas Ahmadi[2]

***Article Info***

Accepted:
27 Apr. 2016

**Keywords**:
Landsat data, Lighvan watershed, Remote sensing, Soil organic carbon

***ABSTRACT***

The current research was directed at Lighvan watershed, northwest of Iran to investigate ETM+ data applicability for the Soil Organic Carbon (SOC) prediction. So, Ground Measurements (GM's) of SOC was carried out in 225 points of study area and ETM+ data were downloaded from NASA's website. Different linear and nonlinear regressions were applied to predict SOC using GM's from whole study area and bare soil only to train the models. Results showed that ETM+ data was impractical for remote sensing of SOC within whole study area due to vegetation effects. Contrary, ETM+ data showed satisfactory accuracy for SOC prediction in bare soils with mean evaluating error (ER) of 18. 34 percent for evaluation stage. A first and second order polynomial between measured SOC and the reflectance of band 1 to 7 of the ETM+ data showed the highest accuracy for SOC prediction with ER of 14 % and $R^2$ of 0.665 for the evaluation stage. Although, ETM+ data application for remote sensing of SOC were restricted by vegetation, it seems that EMT+ data showed enough accuracy for predicting SOC through bare soils.

## INTRODUCTION

Soil Organic Carbon (SOC) is a key parameter for soil mapping, soil properties interpreting, and guiding of organic / chemical fertilizers application (Chen et al. 2000). The SOC affects nearly all soil physical, chemical, and biological properties and dominantly its fertility and productivity (Wu et al. 2009). On the other hand, the SOC decomposition leads to $CO_2$ emission to atmosphere which may result in climate change (Yadav and Malanson 2007). Understanding SOC and its vital rules and spatial and temporal variations are key factors in soil management. Direct measurement and steadily monitoring of SOC at large scales in spite of its necessity may not be feasible. In this regard, several researchers have tried to monitor SOC using remotely sensed data. Most of them have used visual bands to monitor SOC due to the darkness which is usually imparted to the soil by SOC enrichment. Although, Visual bands showed strong correlation with moderate and high amounts of SOC (more than 2%), the correlation was weak in soils with low SOC or also in wet soils (Stephens et al. Unspecified). In the other word, in soils with low SOC, the reflectance of the visual bands are affected by soil moisture rather than SOC. Chen et al. (2000) using aerial images reported a strong linear regression ($R^2$=0.927) between SOC and reflectance from visual red (R), green (G), and blue (B) bands as established below:

$$SOC = \exp(1.71499 - 0.01576R + 0.01281G - 0.0113B) \quad (1)$$

Chen et al. (2005) also carried out an investigation in Georgia using aerial images. Their results also showed a strong exponential relationship ($R^2$=0.930) between SOC and reflectance from visual R, G, and B bands as the following equation:

$$SOC = a_1 \exp(-R/a_4) + a_2 \exp(-G/a_5) + a_3 \exp(-B/a_6) \quad (2)$$

Wu et al. (2009) applied stepwise multivariate regression between digital numbers (DN) of ETM+ bands (ETM1 to ETM6) and SOC reporting significant regression ($R^2$=0.396 and RMSE=0.287) between ln*ETM1* and ln*SOC* ($\ln SOM = a + b \ln ETM1$). The results of

[1-] Department of Soil Science, Faculty of Agriculture, University of Maragheh, Maragheh, Iran
[2-] Department of Soil Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
[3-] Mining Engineering Faculty, Sahand University of Technology, Tabriz, Iran
[4-] Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
*Email: mehdirmti@gmail.com

research performed at Utah University (Stephens et al. Unspecified) showed a reasonably well linear relationship ($R^2$=0.722) between CDI index (extracted from ASTER data) and SOC of surface soil. They defined the CDI as following:

$$CDI = R(2220nm) - R(2270nm) \tag{3}$$

Where, R (2220nm) and R (2270nm) imply reflectance from 2220 and 2270 wavelengths (SWIR), respectively, which are corresponding to the bands 6 and 7 in ASTER data. They reported that although increasing in SOC and soil moisture decreased the reflectance of the different wavelengths, the reflectance of the SWIR was, however, less affected by soil moisture. Contrary, the SOC enhancement led to a nearly uniform reflectance reduction across the all measured spectrum.

Regarding the potential of the remote sensing data to monitor SOC, the current research was carried out to check the applicability of the ETM+ data to predict soil organic carbon (SOC) in Lighvan watershed, Northwest of Iran.

## MATERIALS AND METHODS

### Study area

This research was carried out in the Lighvan watershed, East Azerbaijan, North West of Iran (Fig 1) during spring 2012. The watershed is located at the rangelands of Sahand Mountain in the zone 38N (path=168 and row=34, 37° 43' 07" to 37° 50' 08" N, and 46° 22' 23" to 46° 28' 05" E). The area of the Lighvan watershed is 7,854 hectares and its elevation varies from 3,534m in the uplands to 2,190m at the watershed outlet (Fig 2). The average precipitation of the study area is 320 mm per year. Nearly all parts of the study area have coarse textured soils consisting of loam, sand, sandy clay loam, and sandy loam texture classes (Fig 3). The study area consists four land-uses including bare land, poor pasture, and Irrigated and dry-land farming (Fig 4).

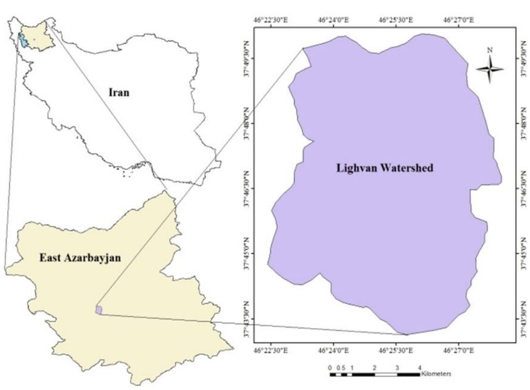


Figure 1. Location of the study area in east Azerbaijan and Iran

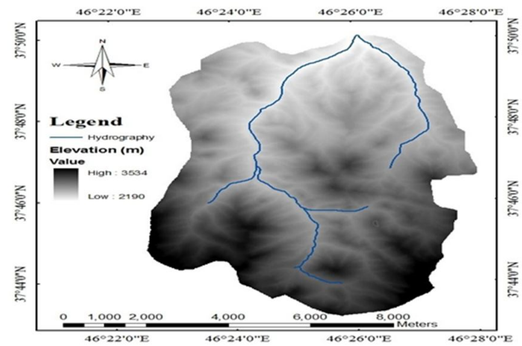


Figure 2. Digital elevation model (DEM) and river network of the study area

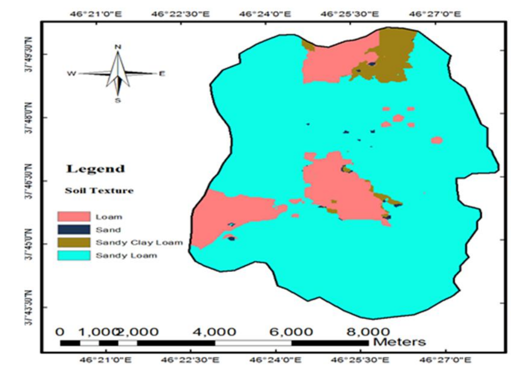


Figure 3. Soil texture map of the study area

### Field and laboratory experiments

Prior to soil sampling, study area was divided into 1 hectare square pixels and then soil samples were taken from pixels. In general, 225 soil samples were taken from 45 cells (five soil samples from each cell). The location of the sampled pixels and the sampling strategy within each pixel is indicated by Figure 5. Soil textures were determined by hydrometer method (Gee and Or, 2002) and soil organic carbon by wet oxidation technique (Nelson and Sommers 1982).

### Landsat data

The images from ETM+ sensor of Landsat7 were used to fulfil the current research. The spatial, spectrum, and temporal resolutions of the ETM+ data are 15 and 30 meters, 8 bands, and 16 days, respectively. In order to carry out this research, ETM+ data of days 13-Jun-2012, 15-July-2012, and 17-Sep-2012 were acquired through USGS website. Finally, the data acquired on 17-Sep-2012 were selected for accuracy analysis due to being less affected by clouds than other days. Prior to use of the satellite data, all needed pre-processing were applied on selected data in order to precise calculations.

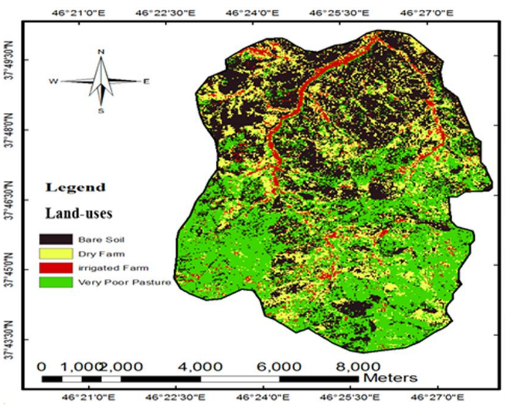


Figure 4. Land-uses map of the study area

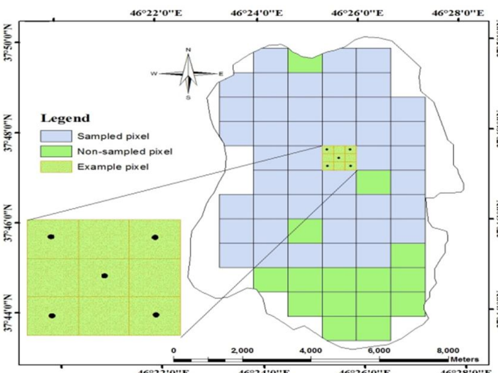


Figure 5. Location of the sampled cells and distribution of sampling points within the study area

**The procedures**

In order to predict soil organic carbon (SOC) through remotely sensed data, two sets of ground measurements (GM's) were applied. Those two sets were included 1) SOC measured in all 225 soil samples taken from four land-uses (bare soil, poor pasture, irrigated and dry-land farms) and 2) SOC measured in soil samples taken from bare soils only.

**Remote sensing of SOC using GM's from whole study area**

In this section, all 225 soil samples taken from whole study area consisting four land-uses (bare soil, poor pasture, irrigated and dry-land farms) were applied to train and evaluate different models. In this regard, several equations reported by different researchers were evaluated for their capability for remote sensing of SOC. On this object, linear and exponential regressions (Chen et al. 2000 and 2005, Wu et al. 2009) were evaluated using GM's of SOC and the reflectance from visual bands (R, G, and B) of ETM+ data:

$$SOC = \exp(a + bR + cG + dB) \tag{4}$$

$$SOC = a_1 \exp\left(\frac{-R}{a_4}\right) + a_2 \exp\left(\frac{-G}{a_5}\right) + a_3 \exp\left(\frac{-B}{a_6}\right) \tag{5}$$

$$\ln(SOC) = a + b \ln(B) \tag{6}$$

Where, SOC implies measured soil organic carbon (%), a, b, c, d, and $a_1$ to $a_6$ imply empirical regression coefficients, and R, G, and B imply the reflectance from red, green and blow visual bands of ETM+ data, respectively. Beside the mentioned equations, a general equation also was applied to predict SOC using reflectance from 7 bands of ETM+ data (bands 1 to 7):

$$SOC = a_1 + a_2 Band1 + a_3 Band2 + \ldots + a_8 Band7 \tag{7}$$

**Remote sensing of SOC using GM's from bare soils only**

Due to the fact that the vegetation may affect the accuracy of the predictions, remote sensing of SOC through bare soils was also examined. In this regard, the four applied regressions form previous section (Eq. 4 to 7) as well as the following polynomials were evaluated using GM's from bare soils only:

a) Second and third order polynomials between measured SOC and reflectance of all bands of ETM+ data

$$SOC = b_1 + b_2 Band1 + b_3 Band1^2 + b_4 Band2 + b_5 Band2^2 + \ldots + b_{15} Band7^2 \tag{8}$$

$$SOC = b_1 + b_2 Band1 + b_3 Band1^2 + b_4 Band1^3 + b_5 Band2 + \ldots + b_{22} Band7^3 \tag{9}$$

b) Linear and exponential regressions between measured SOC and band ratios of ETM+ data:

$$SOC = a + b\left(\frac{Band1}{Band2}\right) + c\left(\frac{Band3}{Band4}\right) + d\left(\frac{Band5}{Band7}\right) \tag{10}$$

$$SOC = a_1 \exp\left(a_4 \times \frac{Band1}{Band2}\right) + a_2 \exp\left(a_5 \times \frac{Band3}{Band4}\right) + a_3 \exp\left(a_6 \times \frac{Band5}{Band7}\right) \tag{11}$$

c) Exponential regression between measured SOC and reflectance from bands 1 to 7 of the ETM+ data:

$$SOC = a_1 \exp(a_2 Band1) + a_3 \exp(a_4 Band2) + \ldots + a_{11} \exp(a_{12} Band7) \tag{12}$$

d) First and second order polynomials between three principle components (PCA1, PCA2, and PCA3) of reflectance of bands 1 to 7 of the ETM+ data and measured SOC:

$$SOC = b_1 + b_2 PAC1 + b_3 PCA2 + b_4 PCA3 \tag{13}$$

$$SOC = b_1 + b_2 PCA1 + b_3 PCA1^2 + b_4 PCA2 + b_5 PCA2^2 + b_6 PCA3 + b_7 PCA3^2 \tag{14}$$

**Model performance**

The applied models behaviour and performance commonly are evaluated and reported through comparisons of observed and simulated variables. In this regard, applied data were divided into two datasets which 60 % and 40 % of data were applied as training (calibration) and validation datasets, respectively. The following criteria beside the determination coefficient ($R^2$) were used to evaluate the SOC models performances:

- Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}\left[X_{t,Obs} - X_{t,Sim}\right]^2}{N}} \quad (15)$$

- Evaluating error

$$ER = \frac{RMSE}{\overline{X_{Obs}}} \times 100 \quad (16)$$

Where, $X_{obs}$ and $X_{Sim}$ imply, respectively, the measured and predicted organic matter at point t with the number of measurements N and the average value of $\overline{X_{Obs}}$ for the measured one. The RMSE, scale-dependent, lies between 0 (perfect fit) and + ∞ (no correlation), and ER, showing error percent, lies between 0 (perfect fit) and + ∞ (no correlation).

## RESULTS AND DISCUSSION

The summery statistics of SOC within the whole study area and bare soils were described in Table 1. The SOC content of study area was laid on between 0.1 and 2 % with average value of 0.85 % (Table 1). Contrary, SOC content of bare soils laid between 0.37 and 1.28 % with average value of 0.78 %.

Table 1. Summery statistics of SOC in whole study area and bare soil

| Variables | Whole study area | Bare soil |
|---|---|---|
| Min | 0.10 | 0.37 |
| Max | 2.05 | 1.28 |
| Mean | 0.85 | 0.78 |
| Standard deviation | 0.40 | 0.17 |

**Predicting SOC using GM's from whole study area**

Different equations as described in previous section were applied for this study. Results showed that the applied equations (Eq. 4 to 7) using remotely sensed data as inputs were impractical for SOC prediction within our study area. The accuracy of the equations were low for both calibration and validation stages, respectively, showing the mean RMSE of 0.361 and 0.413 g/100 g $_{soil}$, mean ER of 42 and 49%, and mean $R^2$ of 0.127 and 0.023. The lower accuracy of the applied equations seems to be due to applying of all GM's from all existence land-uses since Chen et al. (2000 and 2005) and Wu et al. (2009) evaluated the mentioned equations using GM's only from bare soils. So the weak performances of the equations may was expectable due to applying GM's from other land-uses (poor pasture, irrigated, and dry-farm lands) beside bare soils. The general equation (Eq. 7), although, showed slightly better performance with $R^2$=0.181 and 0.055 ($P>0.05$) for both calibration and validation stages, respectively compared to three other applied equations (Eq. 4 to 6), their accuracy was too low to be suggested for remote sensing of SOC in the studied area.

**Predicting SOC using GM's from bare soils**

Results from previous section (Predicting SOC using GM's from whole study area) indicated that the accuracy of the regression models (Eq. 4 to 7) trained by GM's from whole study area consisting bare land, poor pasture, Irrigated and dry-land farming land-uses were too weak. In this regard, varieties of models were applied as described previously. Mathematical expressions of the models have been depicted by Eq. 8 to 14. The models evaluation results are reported by Table 2.

Calibrating the models by applying GM's from bare soil effectively lowered mean ER of the applied models from 42 to 17.1 % for the training stage and from 49.3 to 32.0 % for the validation stage. Furthermore, there was a considerable improvement in determination coefficient ($R^2$) of the measured and predicted SOC showing $R^2$ of 0.331 to 0.127 for training stage and from 0.342 to 0.023 for evaluation stage, respectively.

First and second order polynomial (Eq. 8) showed the highest accuracy (with ER of 13, 14% and $R^2$ of 0.58, 0.62 for the training and evaluation stages, respectively) for remote sensing of the SOC among the all applied equations. Contrary, the proposed equation by Wu et al., (2005), Eq. 6, showed the lowest accuracy (with ER of 20.39 and 21.02 % and $R^2$ of 0.046 and 0.170 for training and evaluation stages, respectively).

It seems that low content (< 2%) of the SOC within the study area is the main reason for the low accuracy of the applied equations especially those require visual bands of the ETM+ data. It is proved that the mentioned bands of the ETM+ data, would show high correlation with high or moderate amount of SOC; the correlation, however, would be low when the SOC decreases below certain level (may be 2%) and/or the soil moisture is high (Stephens et al. unspecified).

Table 2. Evaluation results of the applied regressions for SOC prediction using remote sensing data through bare soil

| Eq. No | Training stage | | | Evaluation stage | | |
|---|---|---|---|---|---|---|
| | RMSE (gr100g$^{-1}$) | ER (%) | $R^2$ | RMSE (gr100g$^{-1}$) | ER (%) | $R^2$ |
| 4 | 0.152 | 19.42 | 0.130 | 0.345 | 21.16 | 0.119 |
| 5 | 0.155 | 19.77 | 0.099 | 0.159 | 20.58 | 0.182 |
| 6 | 0.160 | 20.39 | 0.041 | 0.162 | 21.02 | 0.170 |
| 7 | 0.110 | 14.01 | 0.548 | 0.108 | 14.00 | 0.665 |
| 8 | 0.106 | 13.55 | 0.577 | 0.110 | 14.26 | 0.620 |
| 9 | 0.110 | 14.00 | 0.549 | 0.108 | 14.01 | 0.632 |
| 10 | 0.137 | 17.51 | 0.293 | 0.160 | 20.79 | 0.151 |
| 11 | 0.137 | 17.51 | 0.293 | 0.160 | 20.74 | 0.151 |
| 12 | 0.122 | 15.49 | 0.447 | 0.143 | 18.50 | 0.389 |
| 13 | 0.144 | 18.32 | 0.226 | 0.740 | 96.79 | 0.006 |
| 14 | 0.141 | 17.98 | 0.255 | 0.690 | 90.25 | 0.025 |
| Mean | 0.134 | 17.09 | 0.314 | 0.262 | 32.01 | 0.283 |

RMSE: root mean square error

## CONCLUSION

The remote sensing data were impractical for SOC prediction of our study area considering all existence land-uses in our study area. The accuracy of the applied equations were low showing the mean RMSE of 0.413 g/100 g soil, mean ER of 49%, and mean R2 of 0.023. Contrary, using limited GM's from bare soil considerably led to better prediction of the SOC (with 25 and 31 % lower error for training and evaluation stages, respectively). But generally, authors can claim that Landsat data were found unsuitable for SOC mapping especially in farming lands with considerable vegetation.

## REFERENCES

Chen F. Kissel D.E. West L.T. Adkins W. (2000) Field-scale mapping of surface soil organic carbon using remotely sensed imagery. Soil Science Society of America Journal, 64(2): 746-753.

Chen F. Kissel D.E. West L.T. Rickman D. Luvall J. Adkins W. (2005) Mapping surface soil organic carbon for crop fields with remote sensing. Journal of Soil and Water Conservation, 60(1): 51-57.

Gee G.W. Or D. (2002) 2.4 Particle-size analysis. Methods of Soil Analysis: Physical Methods (Part 4). H. D. Jacob, G. Clarke Topp and Co-Authors, Soil Science Society of America, Inc. Madison, Wisconsin, USA: 255-295.

Nelson D. Sommers L.E. (1982) Total carbon, organic carbon, and organic matter. Methods of soil analysis. Part 2. Chemical and microbiological properties, pp. 539-579.

Stephens S.C. Rasmussen V.P. Ramsey R.D. Whitesides R.E. Searle G.S. Robert L. (Unspecified) Remote Sensing Organic Carbon in Soil, Utah State University.

Wu C. Wu J. Luo Y. Zhang L. DeGloria S.D. (2009) Spatial prediction of soil organic matter content using cokriging with remotely sensed data. Soil Science Society of America Journal, 73(4): 1202-1208.

Yadav V. Malanson G. (2007) Progress in soil organic matter research litter decomposition, modelling, monitoring and sequestration. Progress in Physical Geography, 31(2): 131-154.

**Journal sponsorship**

Azarian Journal of Agriculture is grateful to the ***University of Maragheh*** and its faculty members for their ongoing encouragement, support and assistance.