Kastamonu Uni., Orman Fakültesi Dergisi, 2021, 21(3): 252-267 Research Article
Kastamonu Univ., Journal of Forestry Faculty
Doi:10.17475/kastorman.1049347

Predicting Soil Properties Using Topographic and Climatic Variables

İsmet YENER, Mehmet KÜÇÜK*, Aşkın GÖKTÜRK

Artvin Çoruh University, Faculty of Forestry, Forest Engineering, Artvin, TURKEY
*Corresponding Author: mkck61@artvin.edu.tr

Received Date: 19.11.2020 Accepted Date: 05.09.2021

Abstract

Aim of study: The present study aimed to model soil physical and chemical properties through multiple linear and regression tree techniques.

Area of study: The study area is located between 41,07 – 41,33 N latitude and 41,74 – 42,27 E longitude in Artvin, which is in the Colchis part of the Black Sea Region of Turkey.

Material and methods: The multiple linear regression and regression tree models were used to predict soil properties using topographic and climatic features as independent variables. Besides, the relationships between soil properties and independent variables were determined by Pearson correlation.

Main results: The study results revealed that model accuracy by regression tree generally was higher than those of multiple linear regression. Up to 56% and 59% of the variance in soil properties was accounted for by multiple linear regression and regression tree, respectively. The easting, northing, elevation, and minimum temperature parameters were key drivers of both models. Increasing soil depth significantly increased the pH and reduced the organic carbon, total nitrogen, and carbon/nitrogen ratio.

Highlights: Topographic and climatic variables accounted for Up to 59% and 56% of the variance in soil properties such as texture, pH, organic carbon, total nitrogen, and carbon/nitrogen ratio by regression tree and multiple linear regression techniques. The most influential factors on soil properties were the minimum temperature, latitude, actual
evapotranspiration, mean temperature, distance to the ridge, and radiation index.

Keywords: Forest, Modeling, Multiple Linear Regression, Regression Tree.

Topoğrafik ve Klimatik Değişkenlerden Yararlanarak Toprak Özelliklerinin Tahmin Edilmesi

Öz

Çalışmanın amacı: Bu çalışma ile toprağa ilişkin bazı fiziksel ve kimyasal özelliklerin çoklu doğrusal regresyon ve regresyon ağacı tekniklerinden yararlanılarak modellenmesi amaçlanmıştır.

Çalışma alanı: Çalışma alanı Karadeniz bölgesinin kolşik kesiminde yer alan Artvin ili sınırları içerisinde ve 41,07 – 41,33 K enlemleri ile 41,74 – 42,27 D boylamları arasında bulunmaktadır.

Materyal ve yöntem: Toprak özelliklerinin tahmin edilmesinde çoklu doğrusal regresyon ve regresyon ağacı modelleri kullanılırken, toprak özellikleri ile bağımsız değişkenler arasındaki ilişki ise Pearson korelasyonu ile belirlenmiştir.

Temel sonuçlar: Regresyon ağacı modellerinin doğruluğu çoklu doğrusal regresyon modellerininkine göre daha yüksek bulunmuştur. Toprak özelliklerindeki değişimin en fazla %56 ile %59'luk bir kısmı sırasıyla doğrusal regresyon ve regresyon ağacı modelleri ile açıklanabilmiştir. Her iki model için de en önemli değişkenler boylam, enlem, yükselti ve en düşük sıcaklık olarak belirlenmiştir. Toprak derinliğinin artmasına bağlı olarak pH anlamlı bir şekilde artarken, organik karbon, toplam azot ve karbon/azot oranı azalmıştır.

Araştırma vurguları: Regresyon ağacı modelleri toprak özelliklerindeki değişimi %59'a varan bir oranda, doğrusal regresyon modelleri ise %56'ya varan bir oranda açıklamıştır. Toprak özelliklerini tahminde en belirleyici değişkenler en düşük sıcaklık, enlem, gerçek evapotranspirasyon, ortalama sıcaklık, sırta olan uzaklık ve radyasyon indeksi olarak belirlenmiştir.

Anahtar Kelimeler: Orman, Modelleme, Çoklu Doğrusal Regresyon, Regresyon Ağacı

Introduction

Soil water and soil organic carbon data are essential for simulating climatic, vegetative, and biogeochemical cycles and their responses to change (Henderson et al., 2005). Soil is directly or indirectly related to water quantity and quality, climate change, nutrient cycling, and biodiversity and has become increasingly important. Paralleling its increasing importance, the soil has been under pressure regarding food energy and raw materials (Vogel et al., 2018). The causes of spatial variability in soils are relief, the climate, organisms, the parent material, and time, and how these factors contribute to soil formation in a given period is a very complex process (Hengl et al., 2019). Soil is formed via these factors by Jenny's famous equation (S = f (Cl, O, R, P, T,...)), where Cl is the climate, O are the organisms, R is the relief or the topography, P is the parent material, and T is the time(Van Breemen & Buurman, 2007). Soil property modeling has been needed in recent decades due to difficulties in field studies since that they are expansive and time-consuming methods of measuring/determining soil properties by traditional laboratory techniques, especially on relatively steep slopes. To that end, statistical approaches have been developed to allow fast and accurate processing and evaluate a very complex and large amount of data using instruments and computers. Thus, a new soil science branch was developed, called "pedometrics", the purpose of which is to spatially and temporally model soil data using spatial statistics (Patriche et al., 2011). It is relatively easy to model and evaluate soil properties from topographic factors such as elevation, slope angle, aspect, and distance to the ridge, which affect runoff, drainage soil temperature, erosion, and rate and type of soil formation, in addition to climatic factors such as temperature and precipitation. As in the present study, hillslopes, which have variable topographic and climatic conditions, result in differences in soil properties and require different levels of fertilizers and management systems to optimize production. Many researchers have studied the effect of soil variability on various ecosystem services, such as food provision, timber production, soil retention, water supply, and climate regulation (Adhikari & Hartemink, 2016). The models were developed using the relationships between site factors and soil properties. Multiple linear regression (MLR), the regression tree (RT), the generalized linear model (GLM), and the generalized additive model (GAM) are widely used statistical techniques to predict soil attributes from site factors (Bishop & Minasny, 2006). In addition to MLR, the most-used method, the RT technique has been increasingly used recently because of its predictive power, nonlinear modeling, estimation of qualitative data, and handling of mixed data types (Grunwald, 2016). This study's objective was to develop and evaluate the RT and MLR models to predict soil properties using topographic and climatic variables and then compare the models' predictive performance with performance metrics. The study was performed on the mixed coniferous forests of Artvin, Turkey, in the 2011-2012 summer season.

Materials and Methods

Site Description

The study area is located between 41,07 – 41,33 N latitude and 41,74 – 42,27 E longitude in the Colchis part of the Black Sea Region of Turkey (Figure 1).

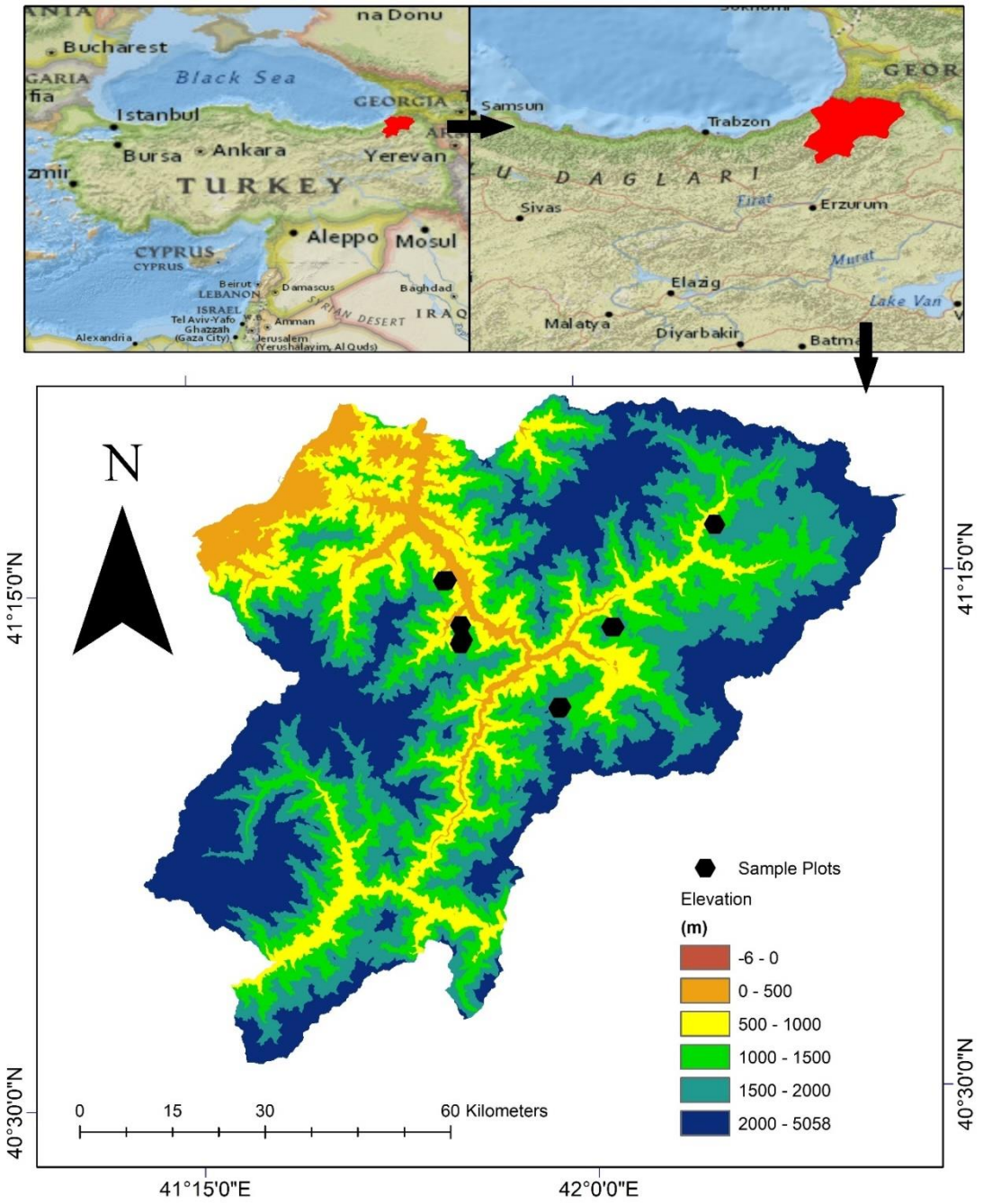

Figure 1. Geographic location of sample plots in Artvin Province (developed using ESRI, 2019; NASA, 2009)

The elevation in the region ranges from 0 to 3930 m above sea level. Steep slopes (>30%) prevail in the province. The soil parent material in the section varies from sedimentary such as marl, conglomerate, and sandstone to upper cretaceous volcanic rocks (Figure 2) mainly found in the study area (Kapur et al., 2017; MTA, 2019). Those rocks include thick piles of andesitic and basaltic lavas, tufts, and agglomerates with a minor amount of trachyandesites and trachytes (Peccerillo & Taylor, 1975) . The soil type is loamy textured acrisols characterized by relatively low pH due to abundant rainfall developed under coniferous trees (Ozcan et al., 2018). The province has placed an importance on its biodiversity, with four plant and nature protection areas

(the Karcal, the Black Sea and Yalnizcam Mountains, and the Coruh Valley), one biosphere reserve area (Camili), three national parks (Karagol-Sahara and Hatila), three nature reserves (Camili-Efeler, Camili-Gorget, and Camburnu) and two nature parks. The province, rich in biodiversity and one of the 34 biodiversity hotspots in the world because of the conservation and protection requirements by the IUCN, accommodating 2727 taxa, 500 (198 endemic and 302 nonendemic) of which are rare and at risk for extinction (Eminagaoglu et al., 2015).

The climate in Artvin varies from arid to semiarid inland, such as in Yusufeli, with 305 mm of annual precipitation, to humid-oceanic in the coastal zone, such as in Hopa with 2230 mm of annual precipitation. The mean annual precipitation and the mean temperature are approximately 1008 mm and 8-12 °C, respectively. The growing season in the region is about six months (from May to October). Although there is a water deficiency in the summer, this deficiency is alleviated by soil moisture utilization (Figure 3).

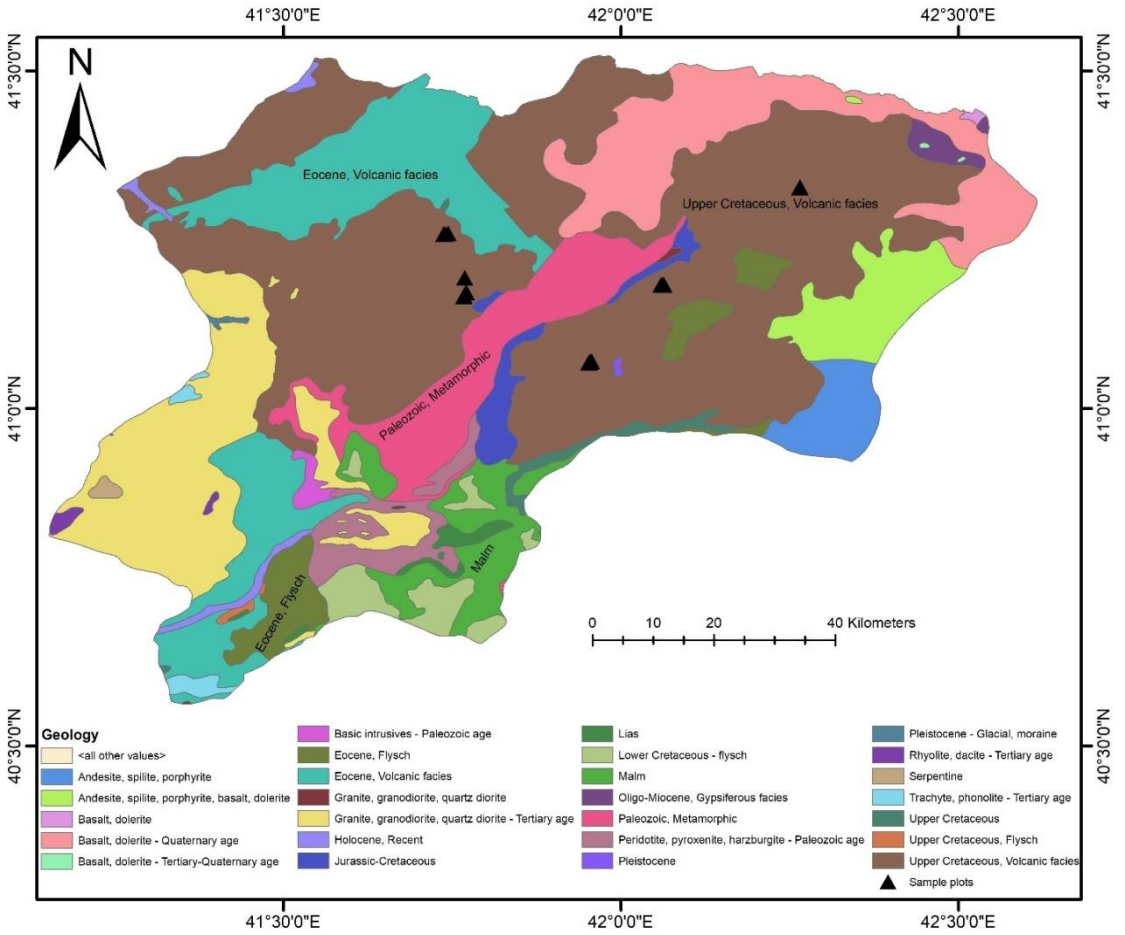

Figure 2. Geological map of Artvin province and the positions of sample plots on it

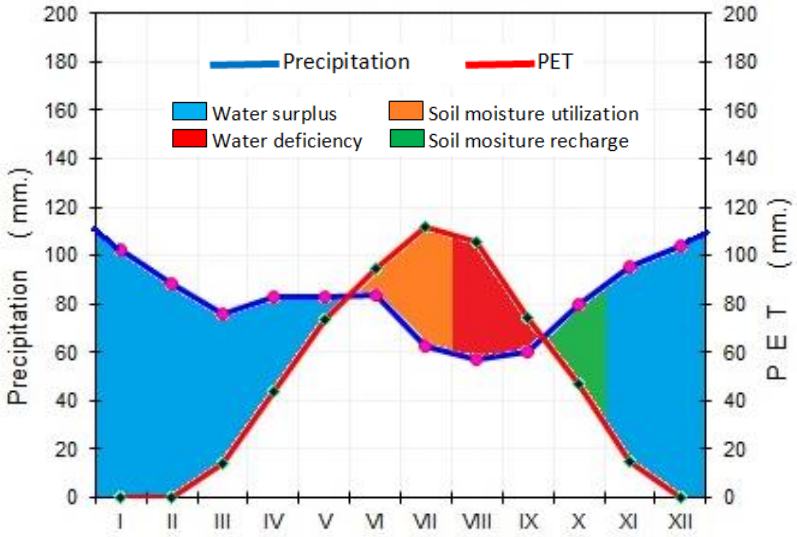

Figure 3. Precipitation and potential evapotranspiration extrapolated to the study site's average elevation (1490 m)

Coniferous species such as *Picea orientalis*, *Pinus sylvestris*, and *Abies nordmanniana* ssp. *nordmanniana*, and deciduous species, such as *Fagus orientalis* and *Carpinus betulus*, are the main species in the province (Eminagaoglu et al., 2015). However, sample plots in the study area contained mixed stands (*Pinus sylvestris-Picea orientalis-Abies nordmanniana* and *Picea orientalis-Pinus sylvestris-Abies nordmanniana*). Along with these mixed stands, *Ilex aquifolium*, *Osmanthus decorus*, *Rhododendron ponticum*, *Viburnum* sp., and *Cornus sanquinea* in Ormanli and Cerattepe and *Pyracantha coccinea*, *Rhododendron luteum*, *Fragaria vesca*, *Vicia faba*, *Celtis australis*, *Rubus* sp., and *Pteridium aquilinum* in Savsat and Ardanuc form the understory vegetation.

Data Collection and Soil Analyses

Sample plots were surveyed during the summer of 2011-2012. Twenty-two sample plots covering a total area of 2500 m^2(50 m × 50 m) were selected. Sample plots were approximately equally distributed on north-facing slopes (NFSs) and south-facing slopes (SFSs). In each sample plot, aspect (ASP), distance to the ridge, where the ridge=0 to the bottom=100 (DTR), elevation (ELEV), and slope (SLP) were determined (Schoeneberger, 2012). For regression analysis, the aspect values (azimuth) were transformed into the radiation index (RADIND), ranging from 0.0 on NNE slopes to 1.0 on SSW slopes using Eq. 1 (Bolton et al., 2018). At each sample plot, coordinates (LAT and LONG) were recorded by handheld GPS with an approximately 5 m accuracy.

$$\text{RADIND} = \frac{\left[1-\cos\left(\left(\frac{\pi}{180}\right)(\text{ASP}-30)\right)\right]}{2} \quad (1)$$

In each sample plot, a soil pit was dug, and four disturbed soil samples were taken at depths of 0-15 cm (SD1), 15-30 cm (SD2), 30-50 (SD3) cm, and >50 cm (SD4). Each soil sample was air-dried and sieved through 2-mm mesh in the laboratory. Then, physical and chemical analyses were performed on 88 disturbed soil samples, with two additional replicates. The soil texture was determined following the hydrometer method. The pH was determined in a 1:2.5 soil distilled water ratio using a glass electrode, and the organic carbon (OC) content was determined by the Walkley and Black wet digestion method. The soil organic matter (OM) was calculated by multiplying the OC content by a coefficient of 1.724. The Kjeldahl digestion, distillation and titration method was used to determine the total nitrogen (TN), and the carbon : nitrogen ratio (C:N) was defined as the ratio of OC to TN (Narwal, 2004). Some climate data, such as the mean annual temperature (Tavg), the mean annual

maximum temperature (Tmax), the mean annual minimum temperature (Tmin), and the annual total precipitation (Prec), were obtained from the meteorological stations closest to the sample plots and extrapolated using 5 °C/km temperature lapse rate (Demircan et al., 2011) and a 450 mm/km increase in precipitation. Using these data, the TavgG, TmaxG, TminG, and PrecG were also computed for the mean annual temperature, the mean annual minimum temperature, the mean annual maximum temperature, and the annual total precipitation in the growing season. The actual evapotranspiration (AET) and the moisture index (TMIND) were also computed based on the Thornthwaite (1948) method.

Data Analyses

The Pearson correlation coefficient was used to examine the relationship between the soil variables and the other factors. Then, to estimate the values regarding various soil properties using topographic and climatic variables, multiple linear regression with the stepwise procedure (MLR) parametric and the regression tree (RT) nonparametric were used. The Pearson correlation, multiple linear regression, and regression tree techniques were performed using R (Team, 2013), SPSS (IBM, 2011), and DTREG v.14 (Sherrod, 2003), respectively. While a one-way ANOVA was used to determine the variation among soil depths, the Tukey HSD test was used for pairwise comparisons. Since the strength of the homogeneity of variances and equal variances were not assumed, the Games-Howell and Welch tests were used. Some globally used predictive performance indices, such as the root mean square error (RMSE), the mean absolute error (MAE), and the correlated Akaike information criteria (AIC_c), in addition to the adjusted coefficient of determination (R^2adj), were selected to evaluate the models using Eqs. 2, 3, 4 and 5.

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

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \quad (3)$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})}{\sum(y_i - \bar{y})}, \; R^2_{adj} = 1 - (1 - R^2)\left[\frac{n-1}{n-(k+1)}\right] \quad (4)$$

$$AICc = (\mathrm{n} * \mathrm{LN}(\mathrm{SSE}/\mathrm{n}) + 2 * \mathrm{k}) + ((2\mathrm{k}\,(\mathrm{k}+1))/(\mathrm{n}-\mathrm{k}-1))) \quad (5)$$

where $\hat{y}$ is the predicted value of y, and $\bar{y}$ is the mean value of y. n is the number of observations, k is the number of estimated parameters, and (k−1) is the number of explanatory variables in the model. After computing the AIC_c scores of the model, the evidence ratio, also called the relative likelihood, was also calculated using Eq. 6, according to Motulsky & Christopoulos (2004).

$$\text{Evidence ratio} = 1/e^{-0.5\Delta AIC_c} \quad (6)$$

where ΔAIC_c shows the difference between the AIC scores.

The most-used predictive performance criterion is the R^2_{adj}, even though some authors claim that it is not a good criterion for comparing different models (Aertsen et al., 2012). Two of the most widely used standards for representing model performance have been the MAE and the RMSE (Chai & Draxler, 2014). On the other hand, Akaike's information criterion (AIC), which determines the most optimal and parsimonious model among competing models by considering model complexity, has recently been developed. The model with the minimum AIC is chosen as the best model. Lower RMSE, MAE, and AIC values and higher R^2_{adj}values indicate better model goodness-of-fit (Pham, 2019).

Results and Discussion

Relationships Between the Physical Soil Properties and Other Variables

The soil texture in the sample plots was loamy sand (1.14%), loamy clay (19.3%), clayey loam (10.2%), sandy loam (30.7%), sandy clay (11.4%), and sandy clayey loam (27.3%). Summary statistics regarding physical soil properties are given in Table 1. The lowest (44%) and the highest (85.5%) sand content in the soils were within depth interval SD3. The sand content did not vary

with soil depth (p=0.065) (Table 1). The positive effect of Y on the sand content within depth intervals SD2 (r=0.51) and SD3 (r=0.44) can be attributed to decreasing precipitation because there is a strong correlation between Y and Prec and PrecG (r=-0.63 and r=-0.68) (Table 1) (Figure 4b, 4c). Increasing rainfall increases the clay content, quickens chemical weathering, and decreases the sand content (Jenny, 1994; Gunal & Ransom, 2006). While the lowest clay content was determined in depth intervals SD1 and SD2 as 5.1%, the highest was within depth interval SD2 with 34.10% (Table 1). According to the one-way ANOVA results (Table 1), the clay content in depth interval SD1 (15.9±1.1%) was significantly lower than in the other soil depth intervals. The relatively low clay content in the topsoil can be attributed to the removal of various materials such as base cations, organic matter, and clay by leaching. The negative effect of Y (r=-0.44) on the clay content, found by correlation analysis, can also be related to decreasing precipitation (Buol et al., 2011).

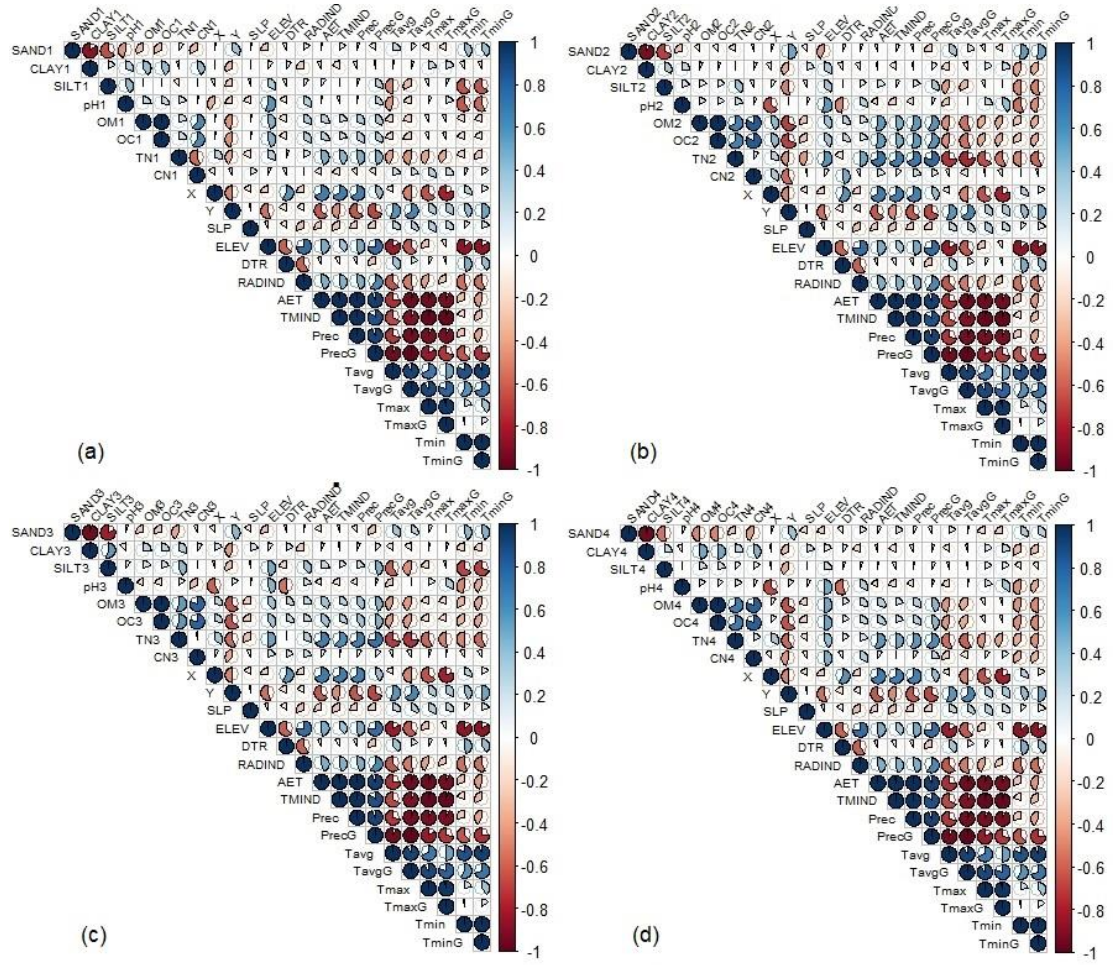

Figure 4. Pearson correlation between variables regarding soil and other ecological factors (A: Soil depth-1 (0-15 cm), B: Soil depth-2 (15-30 cm), C. Soil depth-3 (30-50 cm), and D: Soil depth-4 (>50 cm))

Table 1.Some descriptive statistics regarding soil properties by depth interval

Soil Depth		SAND (%)	CLAY (%)	SILT (%)	pH	OC (%)	TN (%)	C:N
1	$\bar{x}$±SE	68.3±1.5^{a}	15.9±1.1^{a}	15.8±0.8^{a}	5.0±0.1^{a}	2.8±0.2^{a}	0.3±0.0^{a}	11.7±0.7^{a}
	Min.	51.70	5.10	11.80	3.90	1.09	.15	5.65
	Max.	82.50	26.40	22.40	6.61	3.78	.42	16.40
2	$\bar{x}$±SE	61.9±1.9^{a}	21.2±1.4^{b}	16.9±0.8^{a}	5.1±0.1^{b}	1.4±0.1^{b}	0.2±0.0^{b}	8.3±0.4^{b}
	Min.	48.20	5.10	10.90	4.36	.77	.12	5.14
	Max.	80.60	34.10	26.40	6.18	2.59	.23	13.49
3	$\bar{x}$±SE	61.7±2.4^{a}	21.1±1.7^{b}	17.2±1.0^{a}	5.3±0.1^{b}	1.0±0.1^{bc}	0.1±0.0^{bc}	6.9±0.6^{bc}
	Min.	44.00	5.70	8.80	4.46	.46	.09	3.38
	Max.	85.50	33.40	24.70	6.74	2.24	.22	13.79
4	$\bar{x}$±SE	62.4±2.2^{a}	20.8±1.9^{b}	16.7±0.7^{a}	5.4±0.1^{b}	0.7±0.1^{c}	0.1±0.0^{c}	5.3±0.5^{c}
	Min.	45.30	7.10	10.10	4.55	.15	.08	1.94
	Max.	82.00	34.00	22.40	6.88	1.93	.24	9.58
	Sig.	0.065	0.012	0.654	0.044	<0.001	<0.001	<0.001
	F	2.505	4.046	0.543	2.822	48.25	15.738	17.91

* $\bar{x}$±SE: Mean±standard error, Min: Minimum value, Max: Maximum value.
*lower-case letters indicate significant difference by soil depth according to Tukey HSD following one-way ANOVA.

Increasing soil moisture due to precipitation hastens chemical weathering and moves minerals, such as alkaline minerals, deeper into the soil profile. While the minimum silt content was 8.8%, which was in depth interval SD3, the maximum silt content was 26.4%, which was in depth interval SD2. No variation by depth was found for the silt content (Table 1). The silt contents of soils in depth intervals SD1, SD2, and SD3 were negatively correlated with Y and temperatures such as Tavg, TavgG, Tmin, and TminG and positively correlated with ELEV PrecG (Figure 4a, 4b, 4c). The negative effect of temperature on silt content due to increasing elevation can be attributed to increasing precipitation, which directly increases the chemical weathering of parent material and rocks.

Relationships Between the Chemical Soil Properties and the Other Variables

Summary statistics regarding chemical soil properties are given in Table 1. While the minimum pH value, found in depth interval SD1, was 3.90 (very strongly acidic), the maximum value, found in depth interval SD4, was 6.88 (slightly acidic). According to the one-way ANOVA, the pH value in depth interval SD1 was significantly lower than the pH values at the other depths (Table 1).

According to correlation analysis, the pH value of the soils was positively correlated with the ELEV, whose coefficient ranged from r=0.44 to r=0.56 and was negatively correlated with the X (easting), DTR, Tavg, Tmin and TminG (Figure 4a, 4b, 4c, 4d). The positive effect of the ELEV and the negative impacts of temperature, including the Tavg, Tmin, and TminG, on pH can be attributed to accelerated nitrate leaching causing acidification of soils (Kosiba et al., 2018) due to a strong relationship between the elevation and the air temperature, with a 5 °C/km temperature lapse rate. On the other hand, the X and the DTR's adverse effects can be related to soils' increasing water retention capacity. As the DTR (distance to the ridge) increases, the slope decreases. Thus, soil erosion decreases, and clay and organic matter content increase with increasing DTR. All these changes increase the water retention capacity of soils (Kirkham, 2014). Unlike the results in the present study, Zaimes et al. (2017) reported decreased clay content (from 5.42 to 2.99%) in the forested lands of Northern Greece as approaching the streamline. A decrease in the pH with an increase in the X can also be attributed to increased precipitation and moisture index, which remove base cations. In the present study, the correlation

coefficients were 0.66 and 0.67 between X and Prec, and between X and TMIND, respectively (Figure 4a). While the minimum OC content was found within depth interval SD4 with a value of 0.15% (extremely low), the maximum OC content was found within depth interval SD1 with a value of 3.78% (very high) (Table 1). The OC content decreased significantly from the topsoil to the subsoil. Similar results were found by Sariyildiz et al. (2017), who reported 26.9% and 41.7% of decreases in OC in the young and mature fir forest, from topsoil to subsoil. The OC content was positively correlated with the X, ELEV, RADIND, AET, TMIND, Prec, and PrecG values and negatively correlated with the Y, Tavg, TavgG, Tmin, and TminG values. (Figure 4a, 4b, 4c, 4d). Climatic factors such as temperature, precipitation, moisture, and solar radiation affect the decomposition rate of soil organic carbon and the type and growth of plants (Dick & Gregorich, 2004). According to Van't Hoff's temperature rule, an increase in temperature of approximately 10 °C increases the chemical reaction's velocity approximately two to three times.

The decomposition rate of the SOM rises with temperature. Therefore, the SOM decreases with increasing temperature. The negative effect of temperature can be attributed to the increasing decomposition rate, which depends on the rate of chemical reactions. The positive impact of ELEV, TMIND, AET, Prec, and PrecG on soil organic carbon can be related to increasing soil moisture, depending on the factors mentioned above. Miles et al. (2016) reported a positive correlation between organic matter and precipitation (R^2=0.81) and a negative correlation between organic matter and temperature (R^2=0.14) in South African sugarcane soils. For example, Jenny (2012) reported decreasing organic carbon content due to increasing temperature and increasing organic carbon content due to increasing precipitation. He stated that the organic carbon content of soils on Graywackes in New Zealand rose from 4.04 to 13.79 g/m^2 when the mean annual precipitation increased from 345 mm to 1775 mm. In a study done by Selhorst (2011) on home lawn turfgrass development and maintenance in diverse climatic regions of the United States, a positive correlation was found between the SOC and the Prec (R^2=0.13-0.15). Still, after the precipitation reached the threshold (650 mm/year), the relationship changed to negative. While the minimum TN occurred in depth interval SD3 with a value of 0.09%, the maximum TN occurred in depth interval SD1 with a value of 0.42%. The TN content of soils decreased with depth ($p<0.001$, Table 1). TN was positively correlated with ELEV, RADIND, TMIND, AET, Prec, and PrecG and negatively correlated with Y, Tavg, TavgG, Tmax, TmaxG, Tmin, and TminG. The factors related to the OC were also associated with the TN because of the close relationship between OC and TN, ranging from r=0.55 in depth interval SD3 to r=0.68 in depth interval SD4 (Figure 4b, 4c, 4d). The close relationship between the OC and the TN was also stated by Selhorst (2011). They expressed that a 0.1 increase in soil nitrogen caused a 0.99% increase in the SOC. Increasing nitrogen concentration led to a decrease in the SOC decomposition, which increased the mean residence time. The greater the TN, the greater the SOC in forest ecosystems.

The positive effect of the factors mentioned above on the TN can be attributed to increasing soil moisture that improves chemical weathering and decomposition. The negative impact of the Y and temperature can be attributed to the hastening of plant residual decomposition (Buol et al., 2011). Jenny (2012) also reported that the soil organic nitrogen content in North America's Great Plains area decreased approximately 5-fold when the temperature dropped from 20 °C to 0 °C. While the minimum CN (1.94) was found within depth interval SD4, the maximum CN (16.80) was found in depth interval SD1. The CN significantly decreased with depth ($p<0.001$, Table 1). According to the Pearson correlation, the CN was positively correlated with DTR, RADIND, and ELEV and negatively correlated with Y (Figure 4b, 4d). Relief, or distance to the ridge, and slope and aspect, are the factors affecting the local climate, leaching, the transport of materials such as clays, organic matter, carbonates and bases, the

development of soil horizons, and plant nutrition (Kirkham, 2014).

The positive effects of the DTR and elevation can be related to increased soil moisture, which depends on the changing RADIND. In the present study, the DTR was negatively correlated with the RADIND (r=-0.59), which means that as the DTR increased, the aspect turned from SSW to NNE and improved the soil moisture. On the other hand, the positive effect of the ELEV can be attributed to increasing precipitation (r=0.48 and 0.75 for Prec and PrecG, respectively). The climate affected the soil properties in two main ways: directly (the rate and type of soil erosion, weathering, salinization, and desertification) and indirectly (the net primary productivity, hydrological cycles, energy balance, and carbon storage). Due to the effects of precipitation and temperature on soil water availability, the weathering of rocks, transportation, and distribution of weathered material on soil have also been influenced. These climate parameters also affect soil organic carbon accumulation and carbonate leaching. Because of the leaching of base cations in the wet climate, the pH is relatively low, and Al, Mn, and Fe contents are higher than those in the temperate arid climate. The climate also affects the soil biota, which changes the soil organic carbon dynamics (Lal & Stewart, 2018). Effects on soil temperature regimes and soil hydrology due to changes in rainfall and potential evapotranspiration would be expected depending on global climate change. In countries with temperate climate zones, including Turkey, it is anticipated that minor increases in rainfall would be compensated by vegetation and crops through evapotranspiration, which occurs at higher temperatures. Therefore, hydrologic and chemical effects on the soils might be small. The adverse effects on soil organic carbon due to higher temperatures would also be compensated by organic matter from vegetation in better growing circumstances, more elevated CO_2 in the atmosphere, and higher evapotranspiration (Brinkman, 1990).

Modeling of Soil Properties and Selection of the Best Model

Many studies have modeled soil properties so far, using soil-forming factors such as topography and climate. While some researchers (Ziadat, 2005; Hong, 2011; Patriche et al., 2011; Aartsma, 2016; Mason & Sulaeman, 2016; Mosleh et al., 2016; Pinheiro et al., 2018) used only statistical methods such as linear regression, regression tree, generalized additive models, and artificial neural networks, some others (Gutiérrez et al., 2011; Miller et al., 2015; Lin et al., 2016; Liu et al., 2017) used geostatistical techniques such as inverse distance weighting, ordinary kriging, geographically weighted regression, and spline models.

The multiple linear regression (MLR) parametric and the regression tree (RT) nonparametric techniques were used in the present study to model and predict the soil properties in depth intervals SD1, SD2, SD3, and SD4. Model performance metrics such as the R^2_{adj}, RMSE, MAE, and AIC_c are given in Table 2. In the model predictions, three dependent variables in depth intervals SD1 and SD4, and one dependent variable in depth intervals, SD2 and SD3, were not predicted by the MLR. Unlike the MLR, the RT models predicted all of the soil variables (Table 2).

Table 2. Performance indices of models for regression trees and multiple linear regression modeling techniques

	Regression Tree					Multiple Linear Regression					Evidencee
	Variables in	R^2_{adj}	RMSE	MA	AIC_c	Variables in	R^2_{adj}	RMSE	MAE	AIC_c	Ratio (%)
SOIL DEPTH 1											
San	TminG	0.30	4.27	2.62	76.93	No variables were entered into the equation					
Cla	Y, Tmax	0.56	2.53	1.65	56.29	No variables were entered into the equation					
Silt	Tmin	0.58	1.70	1.04	36.30	Tmin, ELEV	0.52	1.78	1.07	40.87	9.8
pH	ELEV	0.39	0.34	0.18	-33.83	Tmin	0.36	0.35	0.19	-32.86	1.6
OC	ELEV	0.42	0.40	0.23	-26.86	Y, RADIND	0.30	0.43	0.28	-21.54	50.4
TN	SLP	0.30	0.04	0.03	-123.45	PrecG	0.18	0.05	0.03	-120.03	5.5
CN	SLP, TmaxG	0.42	1.90	1.23	43.81	No variables were entered into the equation					
SOIL DEPTH 2											
San	TminG	0.38	5.10	3.12	84.69	Tmin	0.25	5.62	3.48	88.99	1.7
Cla	TminG	0.19	4.47	2.53	78.92	No variables were entered into the equation					
Silt	TminG	0.39	2.20	1.30	47.66	TminG	0.23	2.46	1.55	52.59	11.8
pH	AET	0.40	0.29	0.19	-40.89	X, ELEV	0.54	0.25	0.16	-45.63	10.7
OC	SLP	0.46	0.25	0.16	-48.18	Y	0.46	0.25	0.16	-48.01	1.1
TN	SLP	0.53	0.02	0.01	-169.46	PrecG	0.56	0.02	0.01	-171.03	2.2
CN	DTR	0.18	1.29	0.74	24.08	Y, DTR	0.42	1.05	0.65	17.81	22.9
SOIL DEPTH 3											
San	TminG	0.33	6.73	4.07	96.93	TminG	0.17	7.48	4.81	101.60	10.3
Cla	TminG	0.29	4.91	3.13	83.03	Y	0.15	5.37	3.43	87.01	7.3
Silt	TminG, Tmin	0.50	2.35	1.56	53.06	TminG	0.46	2.51	1.68	53.52	1.3
pH	TmaxG	0.33	0.35	0.24	-33.11	X, TavG	0.42	0.32	0.20	-34.96	2.5
OC	SLP	0.46	0.27	0.16	-44.73	Y	0.38	0.29	0.19	-41.79	4.3
TN	AET	0.53	0.02	0.01	-159.56	TavgG	0.52	0.02	0.01	-159.21	1.2
CN	SLP	0.10	1.91	1.17	41.43	No variables were entered into the equation					
SOIL DEPTH 4											
San	TminG	0.27	6.49	4.43	95.33	No variables were entered into the equation					
Cla	TminG	0.23	5.75	3.78	90.01	No variables were entered into the equation					
Silt	ELEV	0.14	2.35	1.41	50.68	No variables were entered into the equation					
pH	AET, Tmax	0.59	0.28	0.17	-40.06	X, PrecG	0.52	0.31	0.19	-36.74	5.3
OC	ELEV, DTR	0.50	0.23	0.14	-49.74	Y	0.39	0.26	0.17	-46.61	4.8
TN	AET	0.44	0.03	0.02	-147.30	PrecG	0.28	0.03	0.02	-141.73	16.2
CN	ELEV, Tavg	0.37	1.41	0.91	30.61	ELEV	0.17	1.66	1.17	35.21	10

The variables in the sand models were Y and X for the RT, and Tmin and TminG for the MLR. The explained variance for the RT was 27 to 48%, with an RMSE from 3.60 to 6.49%; and the explained variance for the MLR was 17 to 25%, with an RMSE from 5.62 to 7.48. The variables in the clay models were X, Y, Tmax, and TminG for RT and Y for MLR. The explained variance was 23 to 56% for the RT, with an RMSE from 2.53 to 5.75; the explained variance was 0.15% for the MLR, with an RMSE of 5.37. No clay models for the MLR were developed except for the SD3 depth interval. The variables in the silt models were X, Tmin, TminG, and ELEV for the RT and Tmin, TminG, and ELEV for the MLR. The explained variance was 14 to 58% for the RT, with an RMSE from 1.70 to 2.35; the explained variance was 23 to 52%, with an RMSE from 1.78 to 2.51%. No MLR silt model was developed for the SD4 depth interval. For soil texture components such as sand, clay, and silt, all RT models with higher R^2_{adj} and lower RMSE, MAE, and AIC_c were superior to the MLR models (Table 2). Y (northing=4555799.5) was the primary split affecting the sand content. The greater the Y, the more sand up to this point. After that point, the X (easting=729685) was determinative (Figure 5a). The other splits predicting sand in depth intervals SD2, SD3, and SD4 were the X. While the primary split for clay is also the Y (4555799.5) before the Tmax (13.75 °C) (Figure 5b), the only split for silt is the Tmin (1.75 °C) (Figure 5c). The X and the TminG predicted the clay content in depth intervals SD2, SD3, and SD4. The X, the TminG, and the ELEV were the other splits predicting silt in depth intervals SD2, SD3, and SD4, respectively. In this study, the explained variance in the soil texture ranged

from 14 to 58% for the RT and from 15 to 52% for the MLR (Table 2). The range for soil texture in other studies showed a wide variation, ranging between 5% and 51% for the MLR and 9% to 69% for the RT (Henderson et al., 2005; Mosleh et al., 2016, Sindayihebura et al., 2017; Pinheiro et al., 2018; Padarian et al., 2019; Zeraatpisheh et al., 2019).

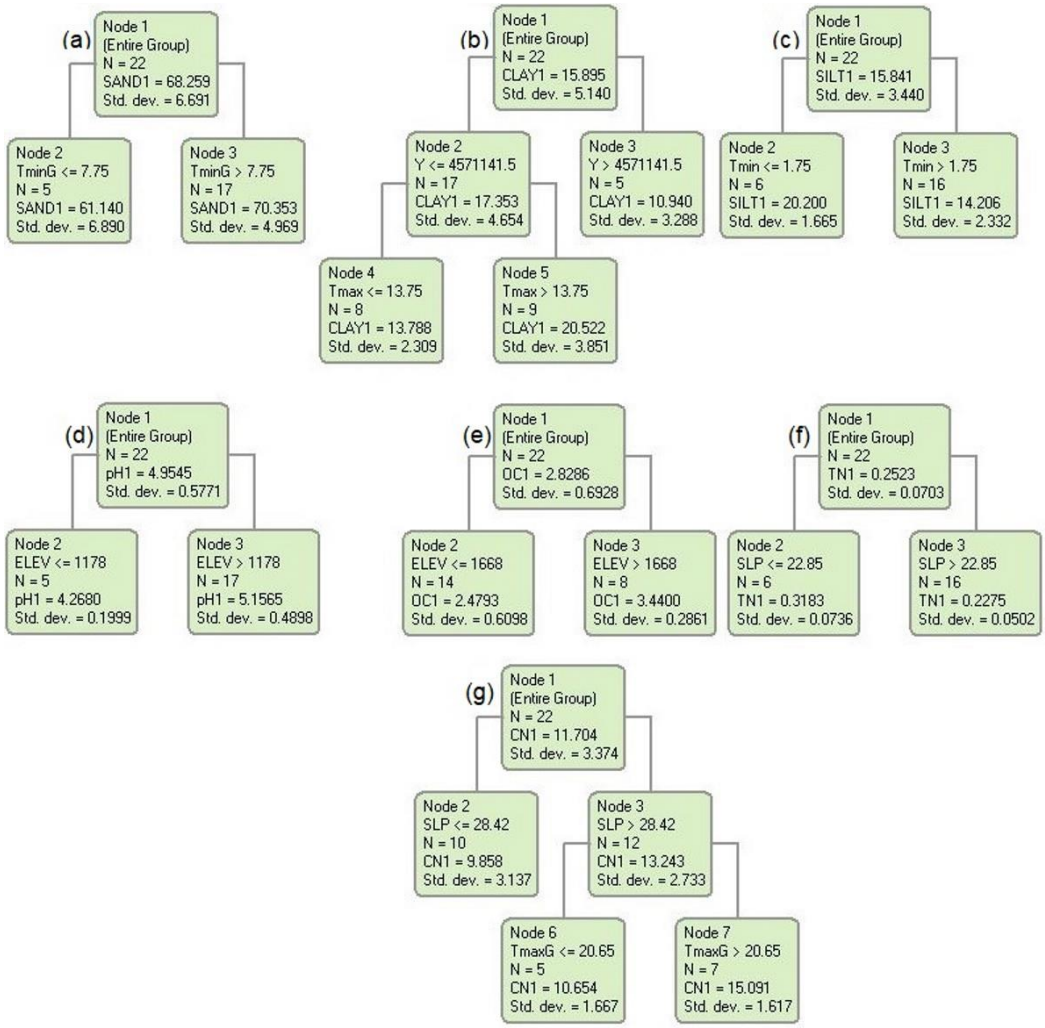

Figure. 5. Regression tree models predicting soil properties in soil depth 1. a) sand, b) clay, c) silt, d) pH, e) organic carbon, f) total nitrogen, and g) carbon/nitrogen)

For example, Pinheiro et al. (2018) developed texture models whose explained variance ranged from 38 to 51% for the MLR and from 48 to 60% for the RT in the Guapu-Macacu watershed, Brazil. Their RT models performed better than their MLR models, as occurred in the present study. Another study (Padarian et al., 2019) used the RT and convolutional neural networks (CNN) to model soil texture. As a result, the RT outperformed the CNN with a higher coefficient of determination ranging from 38 to 42%, with an RMSE from 10.67 to 22.09.

The variables entered into the RT pH models were ELEV, X, Tmax, TmaxG and AET; X, ELEV, Tmin, Tavg, and PrecG were the variables entered into the MLR pH models. The explained variance for the RT was 33 to 59%, with an RMSE from 0.28 to 0.35; the explained variance was 36 to 52%, with an RMSE from 0.25 to 0.42 (Table 2). The pH in the SD1 depth interval was predicted by ELEV (1178 m). The greater the ELEV, the higher the pH (Figure 5d). The other predictors for depth intervals SD2, SD3, and SD4 were X, TmaxG, and AET

and Tmax, respectively. The explained variance determined by the RT in other studies (Henderson et al., 2005; Lourenço et al., 2018; Pahlavan-Rad & Akbarimoghaddam, 2018; Padarian et al., 2019) ranged between 44 and 77%. Pahlavan-Rad & Akbarimoghaddam (2018), for example, modeled pH with an RMSE of 0.20 to 0.45 using the distance to the river, elevation, and slope. The study (Patriche et al., 2011) explained 52% of the variance in the pH with an RMSE of 0.62 by MLR and 48% of it with an RMSE of 0.64 by the RT in the soils of Romania. The predictors for the MLR were Y, elevation, surface ratio, and wetness index, and the predictors for the RT were X, Y, and soil type. Variables in the OC models were Y, ELEV, SLP, and DTR for the RT, and were X, Y, RADIND, and PrecG for MLR. The explained variance was 42 to 50% for the RT, with an RMSE from 0.23 to 0.40, and the explained variance was 30 to 46% for MLR, with an RMSE from 0.25 to 0.43. The OC in depth interval SD1 was also determined using the ELEV (1668 m) (Figure 5e). The other predictors for the OC in depth intervals SD2, SD3, and SD4 were the SLP, SLP and Y, and DTR, respectively. The RT and the MLR were also used in the studies (Mason & Sulaeman, 2016, Mosleh et al., 2016; Zeraatpisheh et al., 2019), and other techniques were used to model soil properties, including terrain attributes, remote sensing, and geology. The R^2_{adj} ranged from 0.04 to 0.36 for the RT and from 0.05 to 0.50 for MLR. While Mason & Sulaeman (2016) found that the RT (R^2±RMSE=0.24±0.69) was better than the MLR (R^2±RMSE=0.05±0.76), Zeraatpisheh et al. (2019) found that MLR (R^2±RMSE=0.36±0.4) was better than the RT (R^2±RMSE=0.50±0.35). The variables in the TN models were X, Y, and SLP for the RT, and TavG and PrecG for MLR. The explained variance for the RT ranged from 30 to 53, with an RMSE from 0.02 to 0.04, and it ranged from 18 to 56 with an RMSE from 0.02 to 0.05 for MLR. The only split for the TN in depth interval SD1 was the SLP (22.85%), which negatively affected the TN (Figure 5f). The other splits for the TN in depth intervals SD2, SD3, and SD4 were the SLP, Y, and X.TheRT was also used to model the TN in soils in some studies (Seiler et al., 2007; Padarian et al., 2019). While Seiler et al. (2007) found that the RT (R^2=0.83 with 0.30 RMSEP) was better than the MLR (R^2=0.72 with 0.49 RMSEP), Padarian et al. (2019) found that the RT (R^2=0.64 with 2.37 RMSE) was better than the other techniques, such as the CNN and partial least squares regression. The variables in the CN models were X, Y, and SLP for the RT, and Y, ELEV, and DTR for MLR. The explained variance for the RT ranged from 19 to 26, with an RMSE from 1.28 to 2.29; the explained variance for MLR ranged from 17 to 42 with an RMSE from 1.05 to 1.66.No variables were entered into the equations for depth intervals SD1 and SD3 for the CN. The CN in depth interval SD1 was also predicted by the SLP (28.42%) (Figure 5g). The other predictors for the TN in depth intervals SD2, SD3, and SD4 were Y, X, and Y. Almost all of the RT models were superior to the MLR models when considering model performance metrics (R^2_{adj}, RMSE, MAE, and AICc), except for the pH, TN, and CN in depth interval SD2 and the pH in depth interval SD3 (Table 2). Considering the evidence ratio, the difference between the AIC_c scores, the RT models were 1.02 to 11.78, 1.62 to 10.66, 1.09 to 50.44, 1.19 to 16.16, 10.01 to 41.25 times more likely correct than the MLR models for the texture, pH, OC, TN, and CN, respectively.

Finally, at local scales, a comparison of the RT technique with MLR to predict soil properties shows that RT outperformed MLR in the ability to increase prediction accuracy.

Conclusion

In this study, soil properties such as the sand, clay, and silt contents, the pH, the organic carbon, the total nitrogen, and the carbon/nitrogen ratio were predicted using soil formation factors such as topography and climate. With this aim, first, the relationships between soil properties and soil formation factors were determined, and then, the most commonly used MLR and RT techniques were employed for developing models. Fifty-six models in total, with twenty-eight from each category, were developed. The most important variables affecting the soil properties were the minimum temperature

and the growing season minimum temperature for the physical soil properties, the latitude (northing), the distance to the ridge, and the minimum temperature for chemical soil properties. Even with very low coefficient determination, the RT models predicted each soil's properties. In contrast, the MLR models predicted each soil's properties except for the sand, clay, and silt contents and the carbon/nitrogen ratio. Considering predictive performance metrics, the RT models with a higher coefficient of determination and a lower RMSE, MAE, and AIC_c were superior to MLR for almost all soil properties.

The IPCC expects that the mean annual temperature will change by 1.6-6.4 °C and the mean annual precipitation will change by at least -20% by 2100, while the atmospheric CO_2 level will increase to 550 ppm. Indeed, we found that the driving climatic factors on the soil properties were temperature, precipitation, actual evapotranspiration, and radiation index. Therefore, climate change may increasingly affect the soil organic carbon and soil water parameters in the future; thus, global climate change is one of the most significant determinants in estimating soil properties.

The techniques that someone with basic statistical knowledge could use for modeling soil properties in the present study are easy to use, user-friendly, and low-cost. The variables to predict soil properties used in the models could also easily be obtained by topographic maps (e.g., DEM) and climate surfaces (e.g., worldclim.org) using GIS. Therefore, the models in the present study may be used mainly in tree species selection for afforestation efforts in the mountainous parts of the regions, requiring considerable cost and labor. However, further studies with large data sets and different tree species are needed to improve these models.

Acknowledgements

This work was supported by Karadeniz Technical University Scientific Research Project Unit, Turkey (Ref. No. 2009.113.001.6).

Ethics Committee Approval

N/A

Peer-review

Externally peer-reviewed.

Author Contributions

Conceptualization: I.Y; Investigation: I.Y, M.K., A.G.; Material and Methodology: I.Y, M.K., A.G. Supervision: I.Y; Visualization: I.Y; Writing-Original Draft: I.Y, M.K., A.G.; Writing-review & Editing: I.Y; Other: All authors have read and agreed to the published version of manuscript.

Conflict of Interest

The authors have no conflicts of interest to declare.

Funding

The authors declared that this study has received no financial support.

References

Aartsma, P. (2016). Climate, A Driving Factor Behind Soil Formation In Proglacial Areas In The European Alps? Master thesis. Master Earth and Environment Soil Geography and Landscape Group, Wageningen University.

Adhikari, K. & Hartemink, A. E. (2016). Linking soils to ecosystem services—A global review. *Geoderma*, 262, 101-111.

Aertsen, W., Kint, V., Muys, B. & Van Orshoven, J. (2012). Effects of scale and scaling in predictive modelling of forest site productivity. *Environmental Modelling & Software*, 31,19-27.

Bishop, T. F. & Minasny, B. (2006). Digital soil-terrain modeling: the predictive potential and uncertainty. *Environmental Soil-landscape Modeling*, 185-213.

Bolton, D. K., White, J. C., Wulder, M. A., Coops, N. C., Hermosilla, T. & Yuan, X. (2018). Updating stand-level forest inventories using airborne laser scanning and Landsat time series data. *International Journal of Applied Earth Observation and Geoinformation*, 66, 174-183.

Brinkman, R. (1990). Resilience Against Climate Change? Pages 51-60, Developments in soil science. *Elsevier*.

Buol, S. W., Southard, R. J., Graham, R. C. & McDaniel, P. A. (2011). Soil genesis and classification. John Wiley & Sons.

Chai, T. & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against

avoiding RMSE in the literature. *Geoscientific model development*, 7, 1247-1250.

Demircan, M., Alan, I. & Sensoy, S. (2011). Increasing resolution of temperature maps by using Geographic Information Systems (GIS) and topography information. 27-29 *in* 5th Atmospheric Science Symposium.

Dick, W. & Gregorich, E. (2004). Developing and maintaining soil organic matter levels. Managing soil quality: Challenges in modern agriculture, 103-120.

Eminagaoglu, O., Akyıldırım Begen, H. & Aksu, G. (2015). Flora and vegetation structure of Artvin (In Turkish). 27-51 *in* O. Eminagaoglu, editor. *Native Plants of Artvin*. Promat, Istanbul.

ESRI. (2019). National geographic world map. http://services.arcgisonline.com/ARCGIS/rest/services/NatGeo_World_Map/MapServer.

Grunwald, S. (2016). Environmental Soil-Landscape Modeling: Geographic Information Technologies and Pedometrics. CRC Press.

Gunal, H. & Ransom, M. D. (2006). Clay illuviation and calcium carbonate accumulation along a precipitation gradient in Kansas. *Catena*, 68, 59-69.

Gutiérrez, Á. G., Contador, F. L., & Schnabel, S. (2011). Modeling soil properties at a regional scale using GIS and Multivariate Adaptive Regression Splines. *Geomorphometry*, 79-82.

Henderson, B. L., Bui, E. N., Moran, C. J. & Simon, D. A. P. (2005). Australia-wide predictions of soil properties using decision trees. *Geoderma*, 124, 383-398.

Hengl, T., Heuvelink, G. & MacMillan, R. (2019). Statistical theory for predictive soil mapping.*in* T. Hengl and R. MacMillan, editors. Predictive Soil Mapping with R. OpenGeoHub Foundation, Wageningen, Netherlands.

Hong, J. (2011). Modeling of soil properties in the Santa Fe River Watershed using exhaustive spatial environmental data. University of Florida.

IBM, C. (2011). IBM SPSS statistics for Windows, version 20.0. New York, IBM Corp 440.

Jenny, H. (1994). Factors of soil formation: a system of quantitative pedology. Courier Corporation.

Jenny, H. (2012). The soil resource: origin and behavior. *Springer Science & Business Media*.

Kapur, S., Akca, E. & Gunal, H. (2017). The Soils of Turkey. *Springer International Publishing*.

Kirkham, M. B. (2014). Principles of soil and plant water relations. Academic Press.

Kosiba, A. M., Schaberg, P. G., Rayback, S. A. & Hawley, G. J. (2018). The surprising recovery of red spruce growth shows links to decreased acid deposition and elevated temperature. *Science of the Total Environment*, 637, 1480-1491.

Lal, R. & Stewart, B. A. (2018). Soil and climate. CRC Press.

Lin, Y., Prentice III, S. E., Tran, T., Bingham, N. L., King, J. Y. & Chadwick, O. A. J. G. R. (2016). Modeling deep soil properties on California grassland hillslopes using LiDAR digital elevation models. *Geoderma Regional*, 7, 67-75.

Liu, W., Zhang, H. R., Yan, D. P. & Wang, S. L. (2017). Adaptive surface modeling of soil properties in complex landforms. *ISPRS International Journal of Geo-Information*, 6(6), 178.

Lourenço, V., Taniguchi, C., Costa, C. A., Toma, R. & Alencar, P. (2018). Using of regression tree for soil organic carbon prediction in the caatinga biome.

Mason, E. & Sulaeman, Y. (2016). Comparison Of Three Models For Predicting The Spatial Distribution Of Soil Organic Carbon In Boalemo Regency, Sulawesi. *Jurnal Ilmu Tanah dan Lingkungan*, 18, 42-48.

Miles, N., Antwerpen, R. V. & Ramburan, S. (2016). Soil organic matter under sugarcane: levels, composition and dynamics. Pages 161-169 *in* Proceedings of the Annual Congress-South African Sugar Technologists' Association. South African Sugar Technologists' Association.

Miller, B. A., Koszinski, S., Wehrhan, M. & Sommer, M. J. G. (2015). Impact of multi-scale predictor selection for modeling soil properties. *Geoderma*, 239, 97-106.

Mosleh, Z., Salehi, M. H., Jafari, A., Borujeni, I. E. & Mehnatkesh, A. (2016). The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environmental Monitoring and Assessment*, 188.

Motulsky, H. & Christopoulos, A. (2004). Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting. Oxford University Press.

MTA. (2019). Geological map of Turkey. General Directorate of Mineral Research and Exploration, Ankara.

Narwal, S. (2004). Research Methods in Plant Sciences: Allelopathy, 1 (Soil Analysis). *Scientific Publishers*.

NASA/METI/AIST/Japan Spacesystems and U.S./Japan ASTER Science Team (2009). ASTER Global Digital Elevation Model [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2021-08-04 from https://doi.org/10.5067/ASTER/ASTGTM.002.

Ozcan, H., Dengiz, O. & Ersahin, S. (2018). Alisols-Acrisols. 207-215 The Soils of Turkey. *Springer*.

Padarian, J., Minasny, B. & McBratney, A. B. (2019). Using deep learning to predict soil properties from regional spectral data. *Geoderma Regional*, 16.

Pahlavan-Rad, M. R. & Akbarimoghaddam, A. (2018). Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). *Catena*, 160, 275-281.

Patriche, C. V., Pîrnău, R. G. & Roşca, B. (2011). Comparing Linear Regression and Regression Trees for Spatial Modelling of Soil Reaction in Dobrovăţ Basin (Eastern Romania). Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca. *Agriculture*, 68.

Peccerillo, A. & Taylor, S. R. (1975). Geochemistry of upper cretaceous volcanic rocks from the pontic chain, northern Turkey. *Bulletin Volcanologique*, 39, 557.

Pham, H. (2019). A New Criterion for Model Selection. *Mathematics*, 7, 1215.

Pinheiro, H. S. K., Carvalho Junior, W. D., Chagas, C. D. S., Anjos, L. H. C. D., & Owens, P. R. (2018). Prediction of topsoil texture through regression trees and multiple linear regressions. *Revista Brasileira de Ciência do Solo*, 42.

Sariyildiz, T., Savaci, G. & Maral, Z. J. (2017). Effect of different land uses (mature and young fir stands-pasture and agriculture sites) on soil organic carbon and total nitrogen stock capacity in Kastamonu Region. *Kastamonu University Journal of Forestry Faculty*, 17, 132-142.

Schoeneberger, P. J. (2012). Field book for describing and sampling soils. National Soil Survey Center, Natural Resources Conservation Service, U.S. Dept. of Agriculture.

Seiler, B., Kneubühler, M., Wolfgramm, B. & Itten, K. (2007). Quantitative assessment of soil parameters in Western Tajikistan using a soil spectral library approach. 451-455 *in* Proceedings of the ISPRS Working Group VII/1 Workshop ISPMSRS'07. *International Society for Photogrammery and Remote Sensing*.

Selhorst, A. L. (2011). Carbon Sequestration By Home Lawn Turfgrass Development and Maintenance in Diverse Climatic Regions of the United States. Ph.D. The Ohio State University.

Sherrod, P. H. (2003). DTREG predictive modeling software. Software available at http://www. dtreg. com.

Sindayihebura, A., Ottoy, S., Dondeyne, S., Van Meirvenne, M. & Van Orshoven, J. (2017). Comparing digital soil mapping techniques for organic carbon and clay content: Case study in Burundi's central plateaus. *Catena*, 156, 161-175.

Team, R. C. (2013). R development core team. *RA Lang Environ Stat Comput*, 55, 275-286.

Thornthwaite, C. W. (1948). An Approach toward a Rational Classification of Climate. *Geographical Review*, 38, 55-94.

Van Breemen, N. & Buurman, P. (2007). Soil Formation. Springer Netherlands.

Vogel, H. J., Bartke, S., Daedlow, K., Helming, K., Kogel-Knabner, I., Lang, B., Rabot, E., Russell, D., Stossel, B., Weller, U., Wiesmeier, M. & Wollschlager, U. (2018). A systemic approach for modeling soil functions. *Soil*, 4, 83-92.

Zaimes, G., Kayiaoglu, K. & Kozanidis, A. (2017). Land-use/vegetation cover and soil erosion impacts on soil properties of hilly slopes in Drama Prefecture of Northern Greece. *Kastamonu University Journal of Forestry Faculty*, 17(3), 427-433.

Zeraatpisheh, M., Ayoubi, S., Jafari, A., Tajik, S. & Finke, P. (2019). Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. *Geoderma*, 338, 445-452.

Ziadat, F. (2005). Analyzing digital terrain attributes to predict soil attributes for a relatively large area. *Soil Science Society of America Journal* 69, 1590-1599.