Black Sea Journal of Agriculture
doi: 10.47115/bsagriculture.1565025

Open Access Journal
e-ISSN: 2618 – 6578

Research Article
Volume 8 - Issue 1: 15-28 / January 2025

DETERMINATION OF ORGANIC CARBON CONTENT OF THE SOILS WITHIN THE GREENHOUSES BUILT ON PYROCLASTIC DEPOSITS IN ISPARTA SETTLEMENT AREA

Sinan DEMİR[1]*, Mehmet Emre ÇAĞ[1]

[1]Isparta University of Applied Sciences, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, 32200, Isparta, Türkiye

Abstract: Soil organic carbon (SOC) is an important indication of soil health and helps to sustain soil fertility. As a result, determining its composition and the factors that influence it is critical for long-term soil nutrient management, especially in controlled conditions such as greenhouses. This study utilizes machine learning to classify SOC content in greenhouses built on pyroclastic deposits in the Isparta region. A dataset of 276 samples and eight variables—clay (%), silt (%), sand (%), soil electrical conductivity (EC), pH, elevation, slope, and aspect—were used to model SOC values. SOC content was classified into five classifications: very low (<0.6%), low (0.6-1.2%), medium (1.2-1.8%), good (1.8-2.3%), and high (>2.3%). In this study, five machine learning models—Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF)—were evaluated using cross-validation to determine their classification accuracy, precision, recall, F-score, and ROC area. Random Forest (RF) and Decision Tree (DT) outperformed the other models, with RF achieving the highest overall accuracy (76.4%), precision (77.3%), and AUC (0.904), followed by DT at 75.4% and AUC of 0.874. This study shows the practicality of machine learning models in categorizing SOC content, highlighting their importance for long-term soil health and fertility control in greenhouse conditions. To improve model efficacy, future studies should include more auxiliary variables, such as soil physical and chemical qualities and lithological data, as well as a wider range of soil types.

Keywords: Soil organic carbon, Machine learning, Greenhouses, Topography, Soil properties, Volcanic materials

***Corresponding author:** Isparta University of Applied Sciences, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, 32200, Isparta, Türkiye
E mail: sinandemir@isparta.edu.tr (S. DEMİR)
Sinan DEMİR https://orcid.org/0000-0002-1119-1186
Mehmet Emre ÇAĞ https://orcid.org/0009-0000-0290-4845

Received: October 11, 2024
Accepted: November 17, 2024
Published: January 15, 2025

Cite as: Demir S, Çağ ME. 2025. Determination of organic carbon content of the soils within the greenhouses built on pyroclastic deposits in Isparta settlement area. BSJ Agri, 8(1): 15-28.

1. Introduction

Soil organic carbon (SOC) is an important component of the global carbon cycle, serving as a crucial carbon store that helps to the management of atmospheric carbon dioxide levels. Within agricultural environments, SOC functions as an active pool that is heavily impacted by human activities. Climate change and soil moisture levels have a significant impact on SOC content, while global warming and population growth highlight the importance of natural (climate, soil parent material, land cover, and topography) and anthropogenic (land use, management, and degradation) factors in SOC dynamics (Hiederer and Köchy, 2011).

The Paris Agreement aims to reduce greenhouse gas emissions by 21% by 2030 when compared to the reference scenario (Genç, 2021). As a result, different projects have been launched globally and in Türkiye to examine SOC levels. Türkiye's SOC status is crucial to promoting sustainable land use and combatting climate change. The importance of estimating and monitoring SOC stocks was stressed at the 12th Conference of the Parties to the United Nations Convention to Combat Desertification in 2015, underlining SOC's role in combating land degradation. The "Türkiye Soil Organic Carbon (CARBON) Project," conducted in collaboration with the General Directorate of Combating Desertification and Erosion (ÇEM) and TÜBİTAK-BİLGEM- Software Technologies Research Institute (YTE), established a high-resolution SOC map using data from 21,061 sampling points. The Random Forest modeling estimated a total carbon stock of 3.51 billion tons in the soil at a depth of 30 cm (ÇEM, 2018). The high-resolution SOC maps and data generated by the CARBON Project aim to enhance agricultural sustainability and develop effective strategies to combat climate change. Several studies have investigated the application of machine learning techniques for predicting SOC, highlighting the effectiveness of various algorithms. For instance, Long Short-Term Memory (LSTM) models demonstrated a high predictive accuracy, with an R^2 value of 0.89 in Southern Xinjiang, China (Wang et al., 2023). Other research has emphasized the importance of spatial SOC distribution, utilizing advanced techniques such as meta-learning stacking to improve predictive performance (Taghizadeh-Mehrjardi et al., 2020). Additionally, recent advancements in digital soil mapping have demonstrated the superiority of models like LSM-ResNet over traditional methods (Zeng et al., 2022). The spatial distribution of SOC has been projected using remote sensing, geographic information systems, and machine learning algorithms (Minasny et al., 2006; Grimm et al., 2008; Minasny et al., 2016; Minasny et al., 2018; Alaboz et al., 2021; Demir and Başayiğit, 2022; Bekana and Mohammed, 2022; Odebiri et al., 2022; Xie et al., 2022; Padarian et al., 2022; Demir, 2024a). In Türkiye, studies focused on SOC determination in greenhouse environments remain limited, necessitating further investigation into the impacts of both anthropogenic and natural factors on SOC levels. Given Türkiye's unique climatic characteristics and agricultural potential, exploring SOC dynamics within greenhouse settings is essential for optimizing soil management practices (TurkStat, 2023).

This study aims to determine the variability of SOC content in greenhouse environments situated on pyroclastic deposits around Isparta. The hypothesis posits that pyroclastic flows and deposits resulting from volcanic activity during the Pliocene and Quaternary periods significantly influence the region's soil composition, leading to notable changes in SOC content. Materials released into the atmosphere during volcanic activity and quickly deposited onto the Earth's surface make up pyroclastic deposits. These deposits can significantly influence soil structure and soil organic carbon (SOC) accumulation when incorporated into surface soils. Pyroclastic materials have a high surface area and fine grain size, which increases their ability to retain water. This helps to build up organic matter and foster plant growth (Elitok et al., 2009; Saputra et al., 2022). The study's primary goal is to develop classification models for SOC levels based on soil characteristics, thereby improving understanding of SOC dynamics and promoting sustainable agricultural practices. To achieve this, soil samples were collected from greenhouses on pyroclastic deposits and analyzed using the Modified Walkley-Black method to determine organic carbon content. Soil pH, electrical conductivity (EC), temperature, texture, and topographic characteristics were also evaluated. Based on the collected soil properties, classification models for SOC levels were constructed using modern technologies and machine learning techniques. This study intends to improve understanding of SOC dynamics in greenhouses built on pyroclastic deposits and to educate sustainable agricultural operations, laying the framework for more efficient and environmentally friendly farming methods.

2. Materials and Methods

2.1. Study Area

The study area is located in the northeastern section of the Central District of Isparta Province, Türkiye, and includes pyroclastic deposits created by volcanic activity near Gölcük Crater Lake (Figure 1). The area is depicted on the M24b3 and M25a4 sheets of the 1:2500 Türkiye Topographic Map. A land survey was undertaken in Deregümü Village and its surroundings, which are 3 kilometers from Isparta's settlement and 1125 meters above sea level.

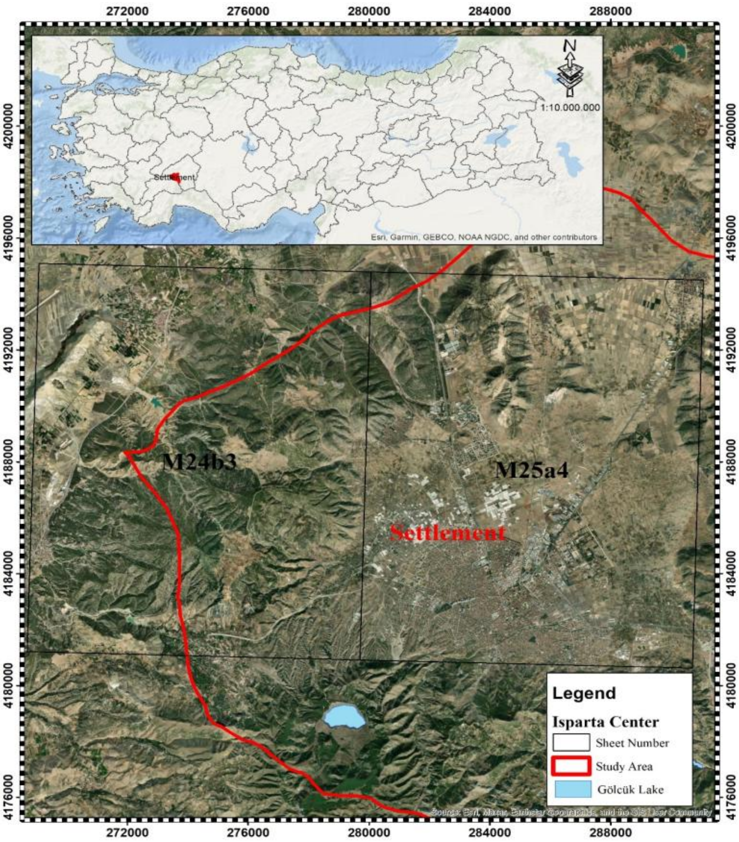

Figure 1. Study area (Isparta Settlement Area) map.

Climate data were obtained from the Turkish State Meteorological Service's Isparta Central Station (MGM). Based on long-term records from 1929 to 2023, the annual average temperature in Isparta was recorded as 12.3 °C, with the highest mean temperature reaching 18.5 °C and the lowest at 6.3 °C. The region receives an average of 7.1 hours of sunshine per day and experiences approximately 99.1 rainy days per year. The highest monthly precipitation was observed in December, totaling 86 mm, while the lowest was recorded in July at 15.5 mm. The maximum recorded temperature in the region was 40.3 °C in August, and the minimum was -21.0 °C in February (MGM, 2024).

According to Corine's 2018 land use data, forest and semi-natural areas account for 64.72% of the total area, covering 50,085.86 hectares. Agricultural land is the second largest land use type, accounting for 29.62% of the total area, covering 22,922.30 hectares. Artificial surfaces account for 5.55% of the total area, covering 4,298.01 hectares. Water bodies account for 0.10% of the total area, covering 80.26 hectares (Corine, 2024). The distribution indicates that the study area predominantly comprises natural and agricultural land uses, ensuring a smooth and clear representation of the region's characteristics.

The 2023 data from the Turkish Statistical Institute (TurkStat, 2023) presents a noteworthy overview of the distribution of agricultural lands and greenhouse activities in the study area. Cereals and other crops make up the largest use of agricultural land, covering 96,911 decares. Next, fallow lands are the second-largest, occupying 26,116 decares. Fruit, beverage, and spice crops follow, covering 21,516 decares. Vegetable cultivation takes up 6,123 decares, while farmers grow ornamental plants on 1,426 decares. In the study area, farmers grow crops in greenhouses that cover 2,506 decares, primarily using plastic for the structures. This indicates that plastic is the most common greenhouse covering material in the region. Farmers do not use other methods like low tunnels, glass greenhouses, or high tunnels (TurkStat, 2023). As a result, greenhouse activities rely solely on plastic structures, limiting the adoption of alternative techniques. These results highlight effective agricultural land use and suggest significant opportunities for expanding greenhouse cultivation in the future.

The Gölcük volcanic region, situated at the apex of the Isparta Angle, represents a geologically complex area influenced by Pliocene and Quaternary volcanic activity. The area is characterized by a combination of autochthonous and allochthonous units that have been intruded by volcanic material and covered with pyroclastic fall and flow deposits. These deposits, consisting of ash, lapilli, and pumice fragments, are associated with the volcanic eruptions of the Gölcük system and have contributed to the formation of surface soils. The pyroclastic material, through its fine-grained structure and high porosity, influences soil physical properties, such as water retention capacity and soil aeration, which are crucial for organic carbon accumulation. Additionally, the pyroclastic deposits are spatially distributed along faults and extensional structures within the Isparta region, further shaping soil development processes (Elitok et al., 2009; Canpolat and Turoğlu, 2019). The interaction of volcanic deposits with tectonic and climatic factors plays an important role in the geochemical genesis of soils in this area, enhancing the study's focus on soil organic carbon dynamics.

2.2. Greenhouses Soil Samples

This study used various cartographic materials in the Remote Sensing and Geographic Information Systems Laboratory of the Department of Soil Science and Plant Nutrition at the Faculty of Agriculture, Isparta University of Applied Sciences to determine the greenhouse areas. These include a 1:2500 scale topographic map, satellite images, major soil group maps, land use capability classes, geological maps, and numerical data (Demir, 2024b). Using Google Earth Pro, 288 greenhouse areas were identified for 2022.

Sample points for this population were calculated using the G-power test, especially the "Means: Equal sample sizes, two groups" test type. This test determines whether the difference in the means of the two groups is statistically significant. In the G-power calculation, the alpha value was set at 5%, and the power (1-beta) was set at 95%, resulting in a 5% margin of error and 95% test power. The effect size was presumably set at 0.5, indicating a medium effect size (Demir et al., 2024).

The calculated sample size was 92, meaning that at least 92 samples were required for analysis. This sample size was expected to detect a medium effect size with 95% power and a 5% margin of error. However, it is important to note that the effect size was estimated hypothetically, and the actual effect size may differ from this assumption. The spatial distribution of the selected sampling points is shown in Figure 2.

Topography characteristics such as elevation, slope, and aspect for each sample point were derived from Türkiye topographic maps. The M24b3 and M25a4 map sheets were digitized, and the corresponding elevation, slope, and aspect values were recorded as attribute data for each point using ArcGIS Pro software (Demir, 2024b). The Corrected Akaike Information Criterion (AICc) was considered, ensuring the ratio of observations to parameters exceeded 40 (Eyduran et al., 2015; Altay, 2022).

Greenhouse farming activities around Isparta are conducted between May and November. During the winter, the coverings are removed, and sampling was performed twice from 92 greenhouses identified for this study. Coordinates of the areas were recorded using GPS, and soil samples were taken from 0-30 cm depth at the end of winter and mid-July (Kacar, 2014). Stratified random sampling was conducted using ArcGIS Pro, identifying 30 points.

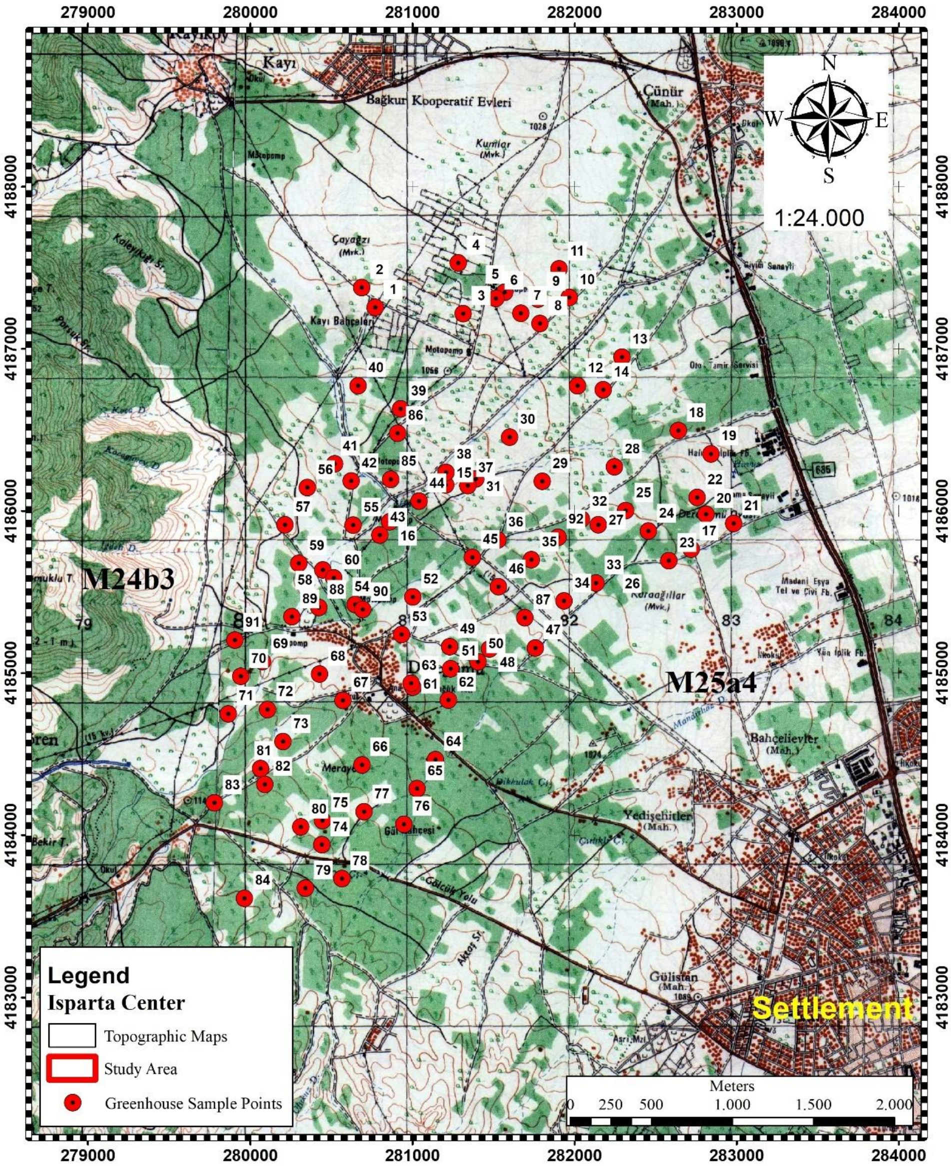

Figure 2. Greenhouse sample points map.

The sampling in two distinct stages was conducted to accurately assess the organic carbon dynamics in the region, particularly considering the practice of greenhouse cultivation in highland areas. During the winter months, production in the greenhouses ceases, and the cover materials are removed, resulting in a significant increase in organic matter inputs. Therefore, the study focuses on the July period, when remote sensing data can effectively capture the status of organic carbon levels, as it reflects the conditions post-cultivation (ÇEM, 2018). By analyzing samples from both periods, this research aims to provide a comprehensive understanding of SOC variations throughout the year, allowing for a better assessment of the seasonal impacts on soil health in the context of highland greenhouse agriculture. Soil sampling from the designated points during the first sampling period was completed between March 19-21, 2023. In the second sampling, additional soil samples were taken from nearby areas with different land uses in July 2023. Soil samples were collected from non-compacted, non-border areas of the plots, placed in polyethylene bags, labeled, and their GPS coordinates recorded. During the land survey, some greenhouses were found to have organic fertilizer added, particularly those growing tomatoes. Carnation greenhouses did not remove their covers in winter. Farmers reported irrigation issues and crop losses due to soil pathogens. Measurements for moisture, temperature, pH, and electrical conductivity (EC) were taken using a KC300 device. Using shading materials in greenhouses to control high temperatures caused changes in light conditions.

2.3. Laboratory Analysis

Soil samples collected from the study area were air dried after being sieved through a 2 mm mesh. The preparation process for the soil samples collected during the first and second sampling periods was completed before analysis. Soil texture analysis was performed using the hydrometer method (Demiralay, 1993). Electrical conductivity (EC) and pH measurements were carried out using a 1:2.5 soil-to-water suspension (Kacar, 2014).

For the determination of SOC air-dried soil samples were sieved through a 500 µm mesh and analyzed using the Modified Walkley-Black method (Kacar, 2014).

These analyses were conducted in the collected soil samples. Organic carbon content was determined on 92 soil samples collected during the first and second sampling periods, with three replicates for each sample. In addition, soil samples collected from adjacent parcels with different land-use types were analyzed with three replicates. Subsequently, calculations were carried out to determine the changes in organic carbon between the two periods, as well as to assess the difference in organic carbon content between the main parcels and adjacent land parcels.

2.4. Statistical Analysis and Modeling

A database was created for the soil samples collected during the first and second sampling periods. Additionally, a dataset was established for the organic carbon content of samples taken from neighboring parcels. The descriptive statistics of the soil samples were evaluated using Minitab 17 software. The normal distribution of the data was checked using the Kolmogorov-Smirnov test (Koşkan et al., 2011; Demir and Başayiğit, 2022). Levene's variance homogeneity test revealed significant differences among the regional types ($P<0.05$); therefore, the Tukey test, significant at $\alpha=0.05$, was employed for post hoc comparisons. This analysis was conducted on 276 different SOC contents and soil organic matter amounts from 92 samples (Demir, 2024a). The organic carbon class range identified through the Turkish Soil Organic Carbon Mapping Project was categorized as very low (<0.6%), low (0.6-1.2%), moderate (1.2-1.8%), good (1.8-2.3%), and high (>2.3%) (Sönmez et al., 2018). Based on these classifications, modeling was performed to classify the observations with the highest accuracy, considering parameters such as sand (%), silt (%), clay (%), pH, electrical conductivity (EC in mmhos/cm), elevation (meters), slope (%), and aspect (°). Classification analyses was conducted using the Weka software (Koçak, 2022). Machine learning models were developed and evaluated, including Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF). In the modeling process, a 10-fold cross-validation method was applied to evaluate the performance of the models more reliably, using the entire dataset. This technique ensures that the reported results are not biased by a particular data split, as each fold is used for training and validation. Model accuracies were determined based on the developed models' General Accuracy, Precision, Sensitivity, F-Score, and ROC area measurements (Demir and Başayiğit, 2022). Cross-validation also helps prevent overfitting by testing the models on multiple subsets of the dataset, providing a more comprehensive performance assessment (Demir et al., 2024).

3. Results

3.1. Soil Organic Carbon

SOC is an important measure of ecosystem health and soil fertility. It includes carbon compounds from plants and other organic materials. These compounds greatly affect the structure of the soil, its ability to hold water, and the nutrients it contains. Proper SOC identification and management are crucial for increasing agricultural production, decreasing erosion, and controlling greenhouse gas emissions. Understanding the link between soil organic carbon and topographic parameters is critical in establishing sustainable soil management and environmental protection methods.

Soil organic matter (SOM) refers to the total organic compounds in the soil and is an important factor in evaluating soil health. It consists of plant leftovers, microbial byproducts, and other organic components. SOC is a key component of SOM, indicating the amount of carbon in organic matter (Demir and Başayiğit, 2021). In general, SOC accounts for 58% of SOM, and this ratio is critical for estimating SOM content. As a result of their relationship with SOM, SOC estimations are critical when assessing soil health and fertility.

Table 1 shows descriptive statistics for soil samples taken in greenhouses. The results from the two sampling periods clearly show variations in SOC and SOM content over time. In the first period, the average SOC value was calculated to be 1.047%, with a standard deviation of 0.516 and a coefficient of variation (CV) of 49.313%. This shows that SOC is very variable, with values that deviate greatly from the mean. In the second period, the average SOC decreased to 0.757%, with a standard deviation of 0.543 and a coefficient of variation of 71.789% (Table 1). The larger CV in this phase indicates a wider range of variability in SOC levels, resulting in a more dispersed distribution than in the previous period. The average SOM content in the first period was calculated to be 1.805%, with a standard deviation of 0.890 and a coefficient of variation (CV) of 49.313%. These numbers suggest considerable variability in SOM, demonstrating that the values deviate significantly from the mean. The average SOM content during the second period was 1.304%, with a standard deviation of 0.936 and a CV of 71.789% (Table 1). The increase in the coefficient of variation during the second phase indicates increased variability and dispersion in SOM data. The variability in SOC and SOM rose significantly across the two periods. This shows that the organic content of the soil samples changed significantly over time, which should be considered when evaluating soil management strategies and greenhouse conditions. The notably high coefficients of variation in the second period highlight the importance of conducting a more in-depth investigation of the effects of soil management practices and environmental conditions on greenhouse soil organic content levels. The seasonal results of the soil samples collected from the greenhouse areas provide important insights into the distribution of SOC and SOM.

Table 1. Descriptive statistics of the dataset of greenhouses

Periods	Variable	*n*	Min.	Mean	Max.	StDev	CoefVar	Skew.	Kurt.
1. Period	SOC, %	276	0.104	1.047	2.650	0.516	49.313	0.400	0.259
	SOM, %		0.179	1.805	4.389	0.890	49.313	0.400	0.259
2. Period	SOC, %	276	0.072	0.757	3.601	0.543	71.789	1.677	1.001
	SOM, %		0.124	1.304	6.208	0.936	71.789	1.677	1.001

It can be observed that the skewness and kurtosis values for both periods align with normal distribution. In the first period, the calculated skewness values for SOC and SOM were both 0.400. These values indicate that the data distribution is symmetric, showing no significant positive or negative skew. The kurtosis values for SOC and SOM were also calculated as 0.259 (Table 1). These results suggest that the data distribution is quite close to normal, with neither a highly peaked nor a flat distribution. This means that the data sets tend to follow a normal distribution, without significant centralization or spreading. In the second period, the skewness and kurtosis values for both SOC and SOM were measured as 1.677. The skewness value of 1.677 indicates a positive skew, meaning the data distribution tends to shift to the right of the mean. Similarly, the kurtosis value of 1.677 suggests that the data distribution is slightly peaked compared to normal distribution, indicating that the data points are somewhat more concentrated around the mean (Table 1). However, these values remain within acceptable limits for normal distribution, indicating that the data set still generally conforms to a normal distribution with no major deviations in terms of skewness or kurtosis.

In conclusion, the skewness and kurtosis values for both SOC and SOM in both periods suggest that the data distribution largely follows a normal pattern. This regularity and consistency in the organic content of the soil samples support the assumption of normal distribution, enhancing the reliability of the statistical analyses. These results indicate that normal distribution assumptions are valid for analytically assessing soil organic carbon content.

During the second period of the study, soil sampling was conducted at points near 30 randomly selected greenhouse areas using a stratified random sampling method. These points were located in different land-use types adjacent to the greenhouses. These areas' SOC and SOM values were compared with those observed under greenhouse conditions. The distribution of land-use types for the sampling points is presented in Table 2. Identifying the various land-use types within the project area is crucial for understanding their impact on SOC and SOM. Each land-use type can have distinct effects on soil health and ecosystem dynamics. For example, vineyards and orchards may enhance soil organic matter content, while fallow land and greenhouse fallow practices play a significant role in soil improvement and sustainable agricultural practices. Other land-use types, such as vegetable and rose gardens, may impact soil fertility and organic matter content in addition to serving aesthetic and commercial purposes.

A detailed analysis of these different land-use types contributes to the development of effective soil management strategies and ecosystem management. Furthermore, the results obtained from the different-sized parcels in agricultural areas provide a broad perspective on the soil organic matter and carbon content, supporting the development of more effective management strategies for these parcels. This evaluation provides valuable insights into the effects of various land-use types on soil properties, which are critical for developing optimized agricultural and ecosystem management practices.

Table 2. The land use type in the side plots of some greenhouse areas

Greenhouse ID Number	Land Use Types
48-49-64-65-70-76-79-80	Vineyard
33-40-52-55-66-83-87-89-90-92	Bare fallow
21-36-88	Greenhouse fallow
5-9	Rose garden
16-26-45	Mixed fruit orchard
7-12	Cherry orchard
1	Walnut orchard
14	Vegetable garden

The descriptive statistics results for the 30 greenhouse sampling points and their adjacent plots have been calculated and are presented in Table 3. In the first period, the average values for SOC and SOM were found to be 1.03% and 1.78%, respectively. The coefficient of variation for SOC and SOM during this period was determined to be 51.92% for both, indicating a wide distribution of the data. The skewness and kurtosis values for both components were measured at 0.47 and 0.39, respectively (Table 3). These values suggest that the data distribution is symmetric and exhibits a tendency close to normal distribution, with data points evenly distributed around the mean. In the second period, the average values for SOC and SOM were calculated as 0.73% and 1.26%, respectively. The coefficients of variation for SOC and SOM were found to be 81.25% for both, indicating a broader variation in the data during this period. The skewness and kurtosis values for SOC were 1.48 and 1.59, and for SOM, they were also 1.48 and 1.59 (Table 3). The positive skewness values indicate a positive trend in the data distribution, with values tending to shift right relative to the mean.

Table 3. Descriptive statistics results in the adjacent plot dataset

Periods	Variable	*n*	Min.	Mean	Max.	StDev	CoefVar	Skew.	Kurt.
1. Period	SOC, %	90	0.104	1.034	2.520	0.537	51.921	0.467	0.392
	SOM, %		0.179	1.783	4.344	0.926	51.921	0.467	0.392
2. Period	SOC, %	90	0.075	0.730	3.367	0.593	81.253	1.480	1.586
	SOM, %		0.129	1.259	5.805	1.023	81.253	1.480	1.586
Adjacent Plot	SOC, %	90	0.047	0.631	2.161	0.467	74.018	1.407	1.725
	SOM, %		0.081	1.088	3.726	0.805	74.020	1.408	1.725

The slightly higher kurtosis values suggest that the data distribution is more peaked than a normal distribution, with data points clustering closer to the mean. For the adjacent plot, the average values for SOC and SOM were calculated as 0.63% and 1.09%, respectively. The coefficients of variation for SOC and SOM in this plot were both 74.02%, indicating a significantly wide distribution of these data. The skewness and kurtosis values for SOC were measured at 1.41 and 1.72, respectively, and for SOM, they were also 1.41 and 1.72 (Table 3). The high values of skewness and kurtosis indicate that the data distribution deviates more from normality, with data points exhibiting a rightward shift and a more peaked distribution compared to a normal distribution.

In conclusion, the skewness and kurtosis values obtained from both periods and the adjacent plot indicate that the data distributions exhibit a certain degree of normal distribution tendency. However, some periods and plots show tendencies of deviation from normality. This suggests that soil organic content varies over time and may exhibit significant differences across different areas. These results provide important insights for soil management and assessment.

The sampling conducted in greenhouse areas resulted in an analysis of SOC content over two different periods. In the first period, the average SOC content was determined to be 1.047% ± 0.516, while in the second period, this value was measured at 0.757% ± 0.543. According to the results of the variance analysis, the difference between the two periods was found to be statistically significant at the 95% confidence level. Based on the Tukey post-hoc test, the SOC content in the first period was classified as "A," and that in the second period as "B" (Table 4). These results indicate that the SOC content in greenhouse areas varies seasonally, and this variation is statistically significant. A higher SOC content was observed in the first period, while this value showed a marked decrease in the second period. This reduction may suggest the influence of seasonal changes, agricultural management strategies employed, or other environmental factors on SOC content. The results indicate that during the winter months in highland conditions, lower temperatures lead to reduced soil biological activity, resulting in decreased soil organic carbon decomposition. However, in open greenhouses, organic matter applications and precipitation contribute to an increase in soil organic carbon levels during this period. In contrast, during the summer months, rising temperatures and heightened soil biological activity are observed, which, coupled with the elevated temperatures within greenhouse conditions, can lead to a reduction in organic carbon content. The determined change in SOC content emphasizes assessing and optimizing greenhouse management and soil improvement measures. Furthermore, it is recommended that more comprehensive studies be conducted to determine whether the seasonal variations are associated with microenvironmental conditions within the greenhouse or changes in agricultural practices. These results could provide a significant foundation for making strategic decisions in greenhouse management and optimizing SOC content.

Table 4. Seasonal SOC content variance analysis results of greenhouse areas

Periods	*n*	Means ± StDev
1. Period	276	1.047±0.516A*
2. Period	276	0.757±0.543B

*= a statistically significant difference exists between the groups (P<0.05).

SOC content in greenhouse areas has been assessed using investigations performed over multiple periods and on adjacent plots. The average SOC content obtained from the adjacent plot is significantly lower (0.631%) than the periodical data gathered from the greenhouse, especially when compared to the SOC content obtained during the first period (1.034%). Furthermore, the SOC content in the second period (0.730%) is consistent with the samples collected from the adjacent plot. However, the initial period's values are significantly higher, indicating a considerable divergence from the SOC contents reported in subsequent periods (Table 5).

Table 5. Variance analysis results of SOC in some greenhouse plots and adjacent plots

Periods	n	Means ± StDev
Adjacent Plot	90	0.631±0.467B*
1. Period	90	1.034±0.537A
2. Period	90	0.730±0.593B

*= a statistically significant difference exists between the groups (P<0.05).

These differences between periods may stem from changes in greenhouse management practices, the effects of agricultural activities, or other environmental conditions. Notably, the high SOC content during the first period suggests an accumulation of organic matter under specific conditions of that period, whereas a significant reduction in SOC content is observed in subsequent periods and the adjacent plot. It is important to consider that the SOC content in the adjacent plot exists at a different level compared to the conditions within the greenhouse and displays similar seasonal variations.

These results underscore the need for assessing the effectiveness of in-greenhouse practices and soil management strategies, as well as evaluating soil characteristics across different plots. Additionally, this data provides a critical foundation for strategic planning aimed at monitoring and improving soil organic carbon levels.

SOC content under greenhouse conditions exhibits significant variations based on periodic analyses. During the first period, the SOC content was determined to be 1.034%, indicating a high accumulation of organic matter in the greenhouse environment. However, in the second period, the SOC content decreased to 0.730%, reflecting a reduction in organic carbon levels. The SOC content obtained from the adjacent plot was found to be 0.631%, which is comparable to the values recorded during the second period. These results demonstrate that the organic carbon content within the greenhouse varies over time. This variability highlights the need to review greenhouse management practices and strategies for organic matter addition.

3.2. Machine Learning Models for Greenhouse SOC

The dataset for classifying SOC content under greenhouse conditions consisted of 276 observations and 8 variables. The variables used in the prediction included clay content (%), silt content (%), sand content (%), electrical conductivity (EC) (mmhos/cm), soil pH value, elevation (meters), slope percentage (%), and aspect direction (degrees). These variables were utilized in the modeling process to predict the soil organic carbon (%OC) content. The descriptive statistics results for these variables are shown in Table 6. The mean clay content was 12.551%, with a Coefficient of Variation of 14.81%, ranging from 9.986% to 15.104%. The clay distribution was relatively symmetrical, indicated by a skewness of -0.03 and a kurtosis of -1.41. The mean silt content was 21.633% (CoefVar = 8.79%), with values ranging from 18.989% to 24.633%. A slight positive skewness of 0.19 and a kurtosis of -0.96 suggested a mild tendency towards higher silt values. Sand content averaged 65.815% (CoefVar = 4.28%), ranging from 62.686% to 71.025%, with a rightward skewness of 0.97 and a kurtosis of -0.30, indicating a distribution close to normal. Electrical conductivity averaged 0.59391 mS/m (CoefVar=25.11%), with values between 0.32 and 0.72 mS/m, demonstrating a negatively skewed distribution (skewness=-1.13) and a kurtosis of -0.66, indicating extreme values. The mean pH level was 7.0773 (CoefVar=8.85%), ranging from 6.0200 to 7.7990, reflecting a negatively skewed distribution (skewness=-0.73) and a kurtosis of -0.92. The average elevation was 1077.0 meters (CoefVar=2.67%), with a range of 1035.0 to 1151.0 meters, showing a mild positive skewness (skewness=0.45) and a kurtosis of -0.73. The mean slope was 2.4096% (CoefVar=53.93%), with values between 0.0000% and 5.6290%, indicating a slight positive skewness (skewness=0.45) and a kurtosis of -0.49. Finally, the average aspect was 115.13° (CoefVar=95.72%), with values ranging from -1.00° to 354.81°, displaying a positive skewness of 1.11 and a kurtosis of -0.09, suggesting a relatively uniform distribution.

In the Türkiye Soil Organic Carbon Mapping Project, the organic carbon content classes were defined as follows: very low (<0.6%) as Class I, low (0.6-1.2%) as Class II, moderate (1.2-1.8%) as Class III, good (1.8-2.3%) as Class IV, and high (>2.3%) as Class V [29] (ÇEM, 2018). The accuracy of the developed machine learning models was evaluated using the Cross-Validation method. This method allows for more reliable testing of each model's performance by repeatedly training and testing on various subsets of the dataset. LR, KNN, SVM, DT, and RF algorithms were systematically tested on the data partitions defined in this process. Cross-Validation helped provide clearer measurements of each model's overall accuracy, precision, sensitivity, F-score, and ROC area, revealing how the models performed across different data subsets.

The advantages and disadvantages of the models can be summarized as follows: LR, while being a simple and fast model, may be limited in capturing complex relationships. KNN can offer high accuracy but may slow down with large datasets.

Table 6. Descriptive statistics results of independent variables dataset (n=276)

Variable	Min.	Mean	Max.	StDev	CoefVar	Skew.	Kurt.
Clay, %	9.99	12.55	15.10	1.86	14.81	-0.03	-1.41
Silt, %	18.99	21.63	24.63	1.90	8.79	0.19	-0.96
Sand, %	62.69	65.82	71.03	2.81	4.28	0.97	-0.30
EC, mmhos/cm	0.32	0.59	0.72	0.15	25.11	-1.13	-0.66
pH,	6.02	7.08	7.80	0.63	8.85	-0.73	-0.92
Elevation, meters	1035.00	1077.00	1151.00	28.80	2.67	0.45	-0.73
Slope, %	0.00	2.41	5.63	1.30	53.93	0.45	-0.49
Aspect, °	-1.00	115.13	354.81	110.20	95.72	1.11	-0.09

SVM achieves high success in classification but may require parameter tuning. DT provides high interpretability but poses a risk of overfitting. RF generally offers high accuracy and minimizes overfitting risk, but the model's explainability can be more challenging. This evaluation process provides a solid foundation for selecting the most suitable approach for predicting organic carbon content by analyzing the basic logic and performance of each model.

When evaluating the performance of the developed classification models, different results were obtained in terms of each model's ability to predict soil organic carbon (%OC) content. The evaluation results are reported here using weighted averages from several categorization algorithms.

The LR model has limitations due to missing metrics such as precision and F-score, making it difficult to fully assess its overall performance. However, it achieved a sensitivity value of 50.7% and an ROC area of 0.595 (Figure 3). These results indicate that the LR model has limited success in predicting soil organic carbon content.

The KNN algorithm reached a sensitivity of 52.5% and an ROC area of 0.611. The performance of KNN could not be fully evaluated due to missing precision and F-score metrics (Figure 3). Nonetheless, KNN showed a slightly higher overall accuracy compared to LR.

The SVM model exhibited lower performance with a sensitivity of 49.6% and an ROC area of 0.552 (Figure 3). The lack of precision and F-score values limited SVM's ability to predict soil organic carbon content.

The DT model demonstrated the highest performance. It achieved a precision of 75.3%, a sensitivity of 75.4%, and an F-score of 75.1. Additionally, the ROC area was determined to be 0.874, with an overall accuracy of 75.4% (Figure 3). These results indicate that the DT model is effective in classifying soil organic carbon content successfully.

The RF model exhibited the highest performance. It attained a precision of 77.3%, a sensitivity of 76.4%, and an F-score of 75.1. The ROC area was determined to be 0.904, with an overall accuracy of 76.4% (Figure 3). RF was observed to be the most successful model in predicting soil organic carbon content.

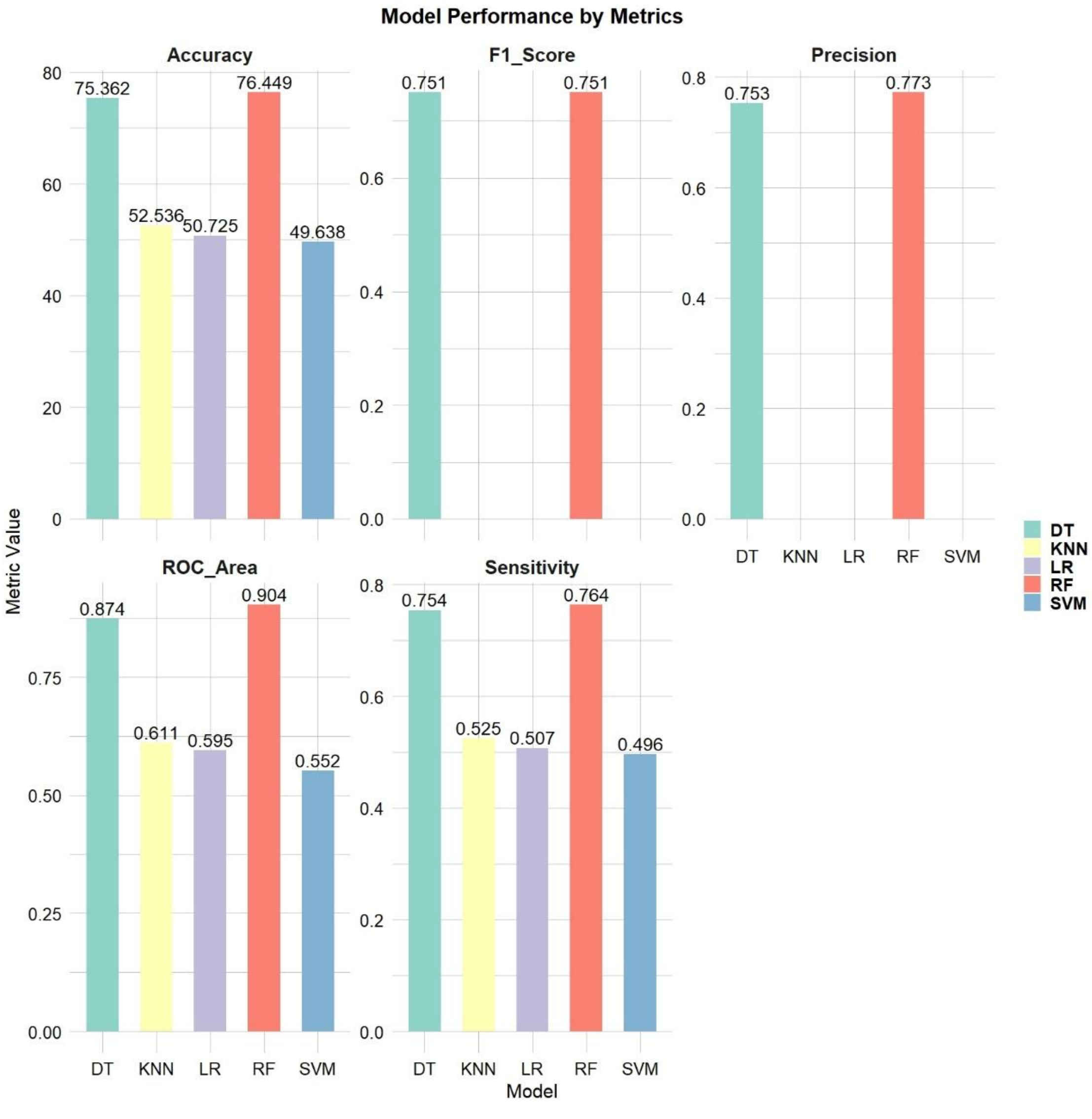

Figure 3. Performance results of SOC machine learning models.

In conclusion, the RF and DT models outperformed other models in classifying soil organic carbon content (Figure 3). These models are regarded as the most suitable methods for effectively predicting soil organic carbon content due to their high accuracy rates and robust performance metrics (Demir and Başayiğit, 2022). Future studies could conduct more detailed analyses on the applicability and improvement of these models.

The model with the highest accuracy in classifying organic carbon in greenhouses is the RF model. The confusion matrix representing the classification performance of the RF model is presented in Figure 4.

The RF model has demonstrated a higher accuracy in classifying organic carbon content compared to other models. The accuracy rate of the RF model is determined to be 76.449%, which reflects the model's overall performance quite successfully.

The confusion matrix illustrates the relationship between the RF model and the class labels:

I. Class (Very Low): In this class, 96 samples were correctly classified, while 20 samples were misclassified. This result indicates that the RF model identifies this class with high accuracy (Figure 4).

II. Class (Low): In this class, 103 samples were correctly classified, whereas 23 samples were misclassified. This indicates that the RF model effectively identifies this class with a high degree of accuracy (Figure 4).

III. Class (Medium): In this class, 7 samples were correctly classified, and 16 samples were misclassified. This suggests that while the model recognizes this class relatively well, it encounters some challenges (Figure 4).

IV. Class (Good): In this class, 4 samples were correctly classified, and 5 samples were misclassified. This result shows that the RF model classifies this class with reasonable accuracy but with some instances of misclassification (Figure 4).

V. Class (High): In this class, 1 sample was correctly classified, and 5 samples were misclassified. This indicates that the model faces challenges in classifying this class (Figure 4).

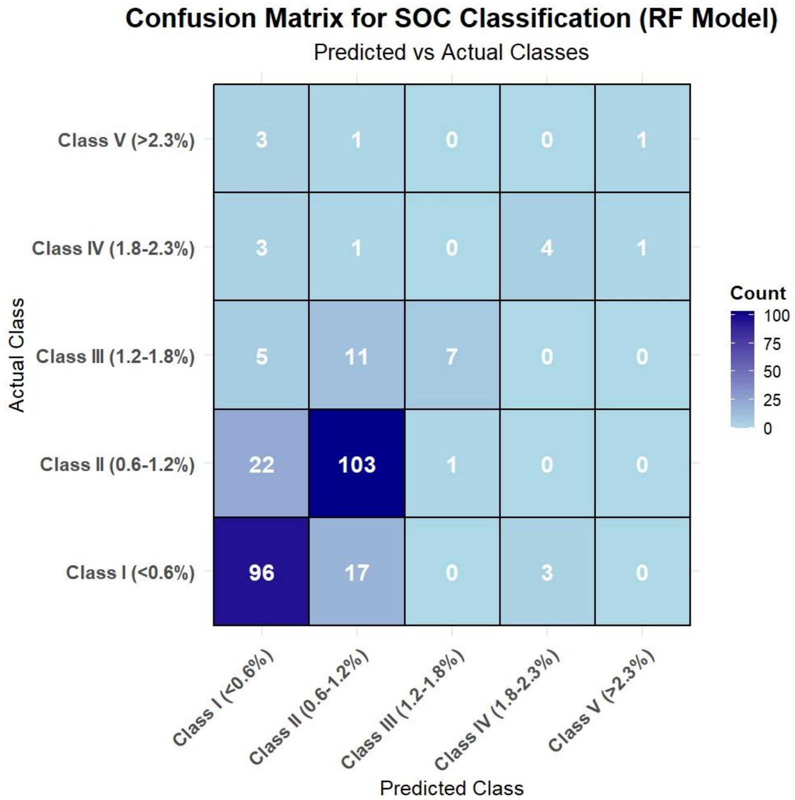

Figure 4. RF model confusion matrix results.

Figure 4 shows the classification results of the RF model, which was the most accurate in recognizing soil organic carbon and indicated overall efficacy. Although the model performs well in certain classes, such as Classes I and II, it is more likely to misclassify others. These results confirm the RF model's effectiveness as a powerful tool for detecting organic carbon soil, but they also highlight the need for enhancements in specific classes. Thus, reviewing model parameters and training data may help improve classification accuracy.

4. Discussion

This study presents substantial results on the performance of SOC content classification using several machine learning models, such as RF, LR, KNN, DT, and SVM. The results show that the RF model has the highest accuracy (76.449%), showing its applicability for SOC classification (John et al., 2020; Zhang et al., 2024). The investigation used a complete dataset that included land and laboratory data as well as topographic features, allowing for high-precision categorical classification of SOCs (Fathizad et al., 2022). A significant asset of this study is the comparison of multiple machine learning methods, which enhances our understanding of their effectiveness in identifying SOC material (John et al., 2020). Regarding the study's advantages, some restrictions were noted. The dataset was geographically restricted, and the sample numbers for several classifications were limited. For instance, Classes III, IV, and V had poorer prediction accuracy, suggesting that it would be challenging to discern between these groups. These results highlight the necessity of more investigation to improve categorization methods and deal with these issues in future studies (Fathizad et al., 2022).

While previous study suggests that the RF model performs well with similar environmental data (Yang et al., 2016; Wang et al., 2018; Fathizad et al., 2022; Loria et al., 2024), it is crucial to note that other algorithms, such as SVM and DT, may perform better under certain scenarios (Bernardini et al., 2024; Agaba et al., 2024). This shows that choosing a suitable model should be context-dependent, taking into consideration the dataset's and study area's unique characteristics. Future research should focus on using larger and more diverse datasets to back up the conclusions of this study. Furthermore, hybrid models and deep learning techniques are proposed to improve model performance (Odebiri et al., 2021; Saporetti et al., 2022; Pouladi et al., 2023; Moharana et al., 2024). It is noteworthy that pyroclastic deposits play a crucial role in soil formation, particularly in the context of SOC development. These deposits, characterized by their rich mineral content and unique physical properties, provide an essential substrate for soil organic matter accumulation (Elitok et al., 2009; Saputra et al., 2022). The influence of pyroclastic materials on nutrient availability and moisture retention can significantly enhance SOC levels, particularly under greenhouse conditions. The study of SOC content under greenhouse conditions on pyroclastic deposits provided valuable insights, emphasizing the necessity for better model parameters and larger datasets for the reliable monitoring of organic carbon levels. While this study did not specifically address economic statistics, it does acknowledge the enormous economic benefits of SOC classification in terms of agricultural output and sustainable farming techniques. Future research should consider these economic implications to develop a more comprehensive understanding of SOC management (Stockmann et al., 2013; Mayer et al., 2020; Derrien et al., 2023).

The conclusions of this study contribute to the development of effective soil management strategies for greenhouse agriculture in the Isparta region. Additionally, the results offer valuable insights into the impact of global warming and climate change on soil organic carbon dynamics, enhancing our understanding of these critical environmental processes.

5. Conclusion

The study successfully proved greenhouse cultivation's effects on organic carbon soil under plateau circumstances in the Isparta region. The information, which included field and laboratory data and topographic features, was evaluated using machine learning techniques, and soil organic carbon concentrations were accurately classified categorically. The results enabled a more precise assessment of soil organic matter levels in greenhouse-growing regions, overcoming the inadequacies of prior approaches and representing a substantial advancement.

The analysis of soil organic carbon content under greenhouse conditions revealed that the RF model was the most accurate, confirming its efficiency in organic carbon classification. However, several misclassifications were observed, particularly in the high and well-represented classes. Based on these results, it is recommended to optimize the model parameters and use larger datasets for more accurate and reliable monitoring of organic carbon in greenhouse conditions. Additionally, a review of greenhouse management practices aimed at increasing organic carbon levels is necessary.

This study contributes to soil management practices for greenhouse horticulture in the region while enhancing our understanding of the impacts of global warming and climate change on soil organic carbon in Türkiye. Further research into the interactions between pyroclastic deposits and organic matter is essential for optimizing soil health and quality. The results offer valuable insights into future land management and sustainable agricultural strategies and serve as a foundation for similar research in other ecosystems.

Author Contributions

The percentages of the authors' contributions are presented below. The author reviewed and approved the final version of the manuscript.

	S.D.	M.E.Ç.
C	90	10
D	100	
S	100	
DCP	50	50
DAI	80	20
L	50	50
W	80	20
CR	60	40
SR	80	20
PM	50	50
FA	50	50

C=Concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

Conflict of Interest

The authors declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

Acknowledgments

This study is based on data derived from the project numbered 1919B012209637, supported by the Research Project Support Programme for Undergraduate Students (2209-A) of the Scientific and Technological Research Council of Türkiye (TUBITAK). I would like to express my sincere gratitude to TUBITAK for their generous funding and support, which enabled this research. I would also like to acknowledge the contributions of individuals who provided invaluable assistance throughout the project.

References

Agaba S, Ferré C, Musetti M, Comolli R. 2024. Mapping soil organic carbon stock and uncertainties in an alpine valley (Northern Italy) using machine learning models. Land, 13(1): 62-78. https://doi.org/10.3390/land13010078

Alaboz P, Dengiz O, Demir S. 2021. Barley yield estimation performed by ANN integrated with the soil quality index modified by biogas waste application. Zemdirbyste, 108(3): 217-226. https://doi.org/10.13080/z-a.2021.108.028

Altay Y. 2022. Prediction of the live weight at breeding age from morphological measurements taken at weaning in indigenous Honamli kids using data mining algorithms. Trop Anim Health Prod, 54(3): 1-12. https://doi.org/10.1007/s11250-022-03174-y

Bekana BT, Mohammed A. 2022. A review of soil organic carbon in Ethiopia's major land use types. Appl Res Sci Technol, 2(1): 25-35. https://doi.org/10.33292/areste.v2i1.21

Bernardini LG, Rosinger C, Bodner G, Keiblinger KM, Izquierdo-Verdiguier E, Spiegel H, Retzlaff CO, Holzinger A. 2024. Learning vs. understanding: When does artificial intelligence outperform process-based modeling in soil organic carbon prediction? N Biotechnol, 81: 20-31. https://doi.org/10.1016/j.nbt.2024.03.001

Canpolat E, Turoğlu H. 2019. The interpretation of the geomorphological development of the volcanic field in the south and southwest of Isparta with lineament and circularity analysis. J Geomorphol Res, 2: 23-36.

Corine. 2024. Ministry of Agriculture and Forestry of the Republic of Türkiye Corine Application. URL: https://corinecbs.tarimorman.gov.tr/ (accessed date: July 21, 2024).

Demir S, Başayiğit L. 2021. The effect of physiographical change on profile development and soil properties. Turk J Agric Res, 8(3): 261-272. https://doi.org/10.19159/tutad.935710

Demir S, Dedeoğlu M, Başayiğit L. 2024. Yield prediction models of organic oil rose farming with agricultural unmanned aerial vehicles (UAVs) images and machine learning algorithms. Remote Sens Appl, 33: 101-131. https://doi.org/10.1016/j.rsase.2023.101131

Demir S. 2024a. Determination of suitable agricultural areas and current land use in Isparta Province, Türkiye, through a linear combination technique and geographic information systems. Environ Dev Sustain, 26: 13455-1349. https://doi.org/10.1007/s10668-023-04359-7

Demir S. 2024b. Prediction of canopy cover for agricultural land classification in Land Parcel Identification System (LPIS) data using Planet-Scope multispectral images: A case of Gelendost District. BSJ Agri, 7(4): 407-417. https://doi.org/10.47115/bsagriculture.1490400

Demir S. Başayiğit L. 2022. Classification of some biochemical properties with J48 classification tree algorithms in hyperspectral data. Veri Bil, 5(2): 20-28.

Demiralay İ. 1993. Physical analysis of soil. The Publications Agricultural Faculty of Atatürk University. No.143. Erzurum, Türkiye, pp: 131.

Derrien D, Barré P, Basile-Doelsch I, Cécillon L, Chabbi A, Crème A, Fontaine S, Henneron L, Janot N, Lashermes G, Quénéa K, Rees F, Dignac MF. 2023. Current controversies on mechanisms controlling soil carbon storage: Implications for interactions with practitioners and policy-makers. Agron Sustain Dev, 43(1): 1-28. https://doi.org/10.1007/s13593-023-00876-x

Elitok Ö, Özgür N, Drüppel K, Dilek Y, Platevoet B, Guillou H. 2009. Origin and geodynamic evolution of late Cenozoic potassium-rich volcanism in the Isparta area, southwestern Turkey. Int Geol Rev, 51(5-6): 454-504. https://doi.org/10.1080/00206810902951411

Eyduran SP, Akin M, Ercisli S, Eyduran E, Maghradze D. 2015. Sugars, organic acids, and phenolic compounds of ancient grape cultivars (Vitis vinifera L.) from Iğdır province of Eastern Türkiye. Biol Res, 48: 1-8. https://doi.org/10.1186/0717-6287-48-2

Fathizad H, Taghizadeh-Mehrjardi R, Hakimzadeh Ardakani MA, Zeraatpisheh M, Heung B, Scholten T. 2022. Spatiotemporal assessment of soil organic carbon change using machine learning in arid regions. Agronomy, 12(3): 615-628. https://doi.org/10.3390/agronomy12030628

Genç C. 2021. Türkiye's obligations under the Paris Climate Agreement and the impact of climate change on these obligations. MSc Thesis, İskenderun Technical University, Graduate Institute, Department of Aquaculture, Hatay, Türkiye, pp: 133.

General Directorate of Combating Desertification and Erosion (ÇEM). 2018. Soil organic carbon project, technical summary.

URL: https://kutuphane.tarimorman.gov.tr/vufind/Record/1178980 (accessed date: July 21, 2024).

Grimm R, Behrens T, Märker M, Elsenbeer H 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. Geoderma, 146(1-2): 102-113. https://doi.org/10.1016/j.geoderma.2008.05.008

Hiederer R, Köchy M. 2011. Global soil organic carbon estimates and the harmonized world soil database, EUR 25225 EN, Publications Office of the European Union, Brussels, Belgium, pp: 79.

John K, Abraham Isong I, Michael Kebonye N, Okon Ayito E, Chapman Agyeman P, Marcus Afu S. 2020. Using machine learning algorithms to estimate soil organic carbon variability with environmental variables and soil nutrient indicators in an alluvial soil. Land, 9(12): 467-487. https://doi.org/10.3390/land9120487

Kacar B. 2014. Soil analysis. Nobel Akademik Yayıncılık, Ankara, Türkiye, pp: 467.

Koçak H. 2022. Performance comparison of classification algorithms in rainfall prediction. J Intell Syst: Theory Appl, 5(1): 16-26. https://doi.org/10.38016/jista.979285

Koşkan Ö, Önder EG, Şen N. 2011. Use of canonical correlation for estimating relationship between variable sets. J Inst Sci Technol, 1(2): 117-123.

Loria N, Lal R, Chandra R. 2024. Handheld in situ methods for soil organic carbon assessment. Sustain, 16(13): 55-92. https://doi.org/10.3390/su16135592

Mayer M, Prescott CE, Abaker WE, Augusto L, Cécillon L, Ferreira GWD, James J, Jandl R, Katzensteiner K, Laclau JP, Laganière J, Nouvellon Y, Paré D, Stanturf J A, Vanguelova EI, Vesterdal L. 2020. Tamm review: Influence of forest management activities on soil organic carbon stocks: A knowledge synthesis. For Ecol Manage, 466: 118100-118127. https://doi.org/10.1016/j.foreco.2020.118127

MGM. 2024. Türkiye State Meteorological Service: Isparta climate data. Available from: https://www.mgm.gov.tr/Veridegerlendirme/il-ve-ilceler-istatistik.aspx?m=ISPARTA (accessed date: 21 Jul 2024).

Minasny B, McBratney AB, Mendonça-Santos MDL, Odeh IOA, Guyon B. 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. Soil Res, 44(3): 233-244. https://doi.org/10.1071/SR05136

Minasny B, Setiawan BI, Arif C, Saptomo SK, Chadirin Y. 2016. Digital mapping for cost-effective and accurate prediction of the depth and carbon stocks in Indonesian peatlands. Geoderma, 272: 20-31. https://doi.org/10.1016/j.geoderma.2016.02.026

Minasny B, Setiawan BI, Saptomo SK, McBratney AB. 2018. Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands. Geoderma, 313: 25-40. https://doi.org/10.1016/j.geoderma.2017.10.018

Moharana PC, Yadav B, Malav LC, Kumar S, Meena RL, Nogiya M, Biswas H, Patil NG. 2024. Regional prediction of soil organic carbon dynamics for intensive farmland in the hot arid climate of India using the machine learning model. Environ Earth Sci, 83: 529. https://doi.org/10.1007/s12665-024-11834-5

Odebiri O, Mutanga O, Odindi J, Naicker R. 2022. Modelling soil organic carbon stock distribution across different land-uses in South Africa: A remote sensing and deep learning approach. ISPRS J Photog Remote Sens, 188: 351-362. https://doi.org/10.1016/j.isprsjprs.2022.04.026

Odebiri O, Odindi J, Mutanga O. 2021. Basic and deep learning models in remote sensing of soil organic carbon estimation: A brief review. Int J Appl Earth Obs Geoinf, 102: 102389. https://doi.org/10.1016/j.jag.2021.102389

Padarian J, Stockmann U, Minasny B, McBratney AB. 2022. Monitoring changes in global soil organic carbon stocks from space. Remote Sens Environ, 281: 113260. https://doi.org/10.1016/j.jag.2021.102389

Pouladi N, Gholizadeh A, Khosravi V, Borůvka L. 2023. Digital mapping of soil organic carbon using remote sensing data: A systematic review. Catena, 232: 107409. https://doi.org/10.1016/j.catena.2023.107409

Saporetti CM, Fonseca DL, Oliveira LC, Pereira E, Goliatt L. 2022. Hybrid machine learning models for estimating total organic carbon from mineral constituents in core samples of shale gas fields. Mar Pet Geol, 143: 105783. https://doi.org/10.1016/j.marpetgeo.2022.105783

Saputra DD, Sari RR, Hairiah K, Widianto Suprayogo D, van Noordwijk M. 2022. Recovery after volcanic ash deposition: Vegetation effects on soil organic carbon, soil structure, and infiltration rates. Plant Soil, 474: 163-179. https://doi.org/10.1007/s11104-021-05226-5

Sönmez B, Özbahçe A, Akgül S, Keçeci M. 2018. Development of a geographic database for the fertility and organic carbon (SOC) content of Turkish soils. Final Project Report (TAGEM/TSKAD/11/A13/P03). Soil, Fertilizer, and Water Resources Central Research Institute. General Directorate of Agricultural Research and Policies, Ankara, Türkiye. URL: https://arastirma.tarimorman.gov.tr/toprakgubre/Belgeler/2018%20Y%C4%B1l%C4%B1%20Proje%20Raporlar%C4%B1/T%C3%BCrkiye%20Topraklar%C4%B1n%C4%B1n%20Baz%C4%B1%20Verimlilik%20ve%20Organik%20Karbon%20(TOK)%20%C4%B0%C3%A7eri%C4%9Finin%20Co%C4%9Frafi%20Veritaban%C4%B1n%C4%B1n%20Olu%C5%9Fturulmas%C4%B1.pdf (accessed date: July 21, 2024).

Stockmann U, Adams MA, Crawford JW, Field DJ, Henakaarchchi N, Jenkins M, Zimmermann M. 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric Ecosyst Environ, 164: 80-99. https://doi.org/10.1016/j.agee.2012.10.001

Taghizadeh-Mehrjardi R, Schmidt K, Amirian-Chakan A, Rentschler T, Zeraatpisheh M, Sarmadian F, Valavi R, Davatgar N, Behrens T, Scholten T. 2020. Improving the spatial prediction of soil organic carbon content in two contrasting climatic regions by stacking machine learning models and rescanning covariate space. Remote Sens, 12(7): 1095. https://doi.org/10.3390/rs12071095

TurkStat. 2023. Türkiye Statistical Institute: Isparta central crop production statistics. URL: https://biruni.tuik.gov.tr/medas/?locale=tr (accessed date: July 21, 2024).

Wang B, Waters C, Orgill S, Cowie A, Clark A, Liu DL, Sides T. 2018. Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. Ecol Indicat, 88: 425-438. https://doi.org/10.1016/j.ecolind.2018.01.049

Wang Y, Chen S, Hong Y, Hu B, Peng J, Shi Z. 2023. A comparison of multiple deep learning methods for predicting soil organic carbon in Southern Xinjiang, China. Comput Electron Agric, 212: 108067. https://doi.org/10.1016/j.compag.2023.108067

Xie B, Ding J, Ge X, Li X, Han L, Wang Z. 2022. Estimation of soil organic carbon content in the Ebinur Lake wetland, Xinjiang, China, based on multisource remote sensing data and ensemble learning algorithms. Sensors, 22(7): 2685.

https://doi.org/10.1016/j.compag.2023.108067
Yang RM, Zhang GL, Liu F, Lu YY, Yang F, Yang F, Li DC. 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. Ecol Indicat, 60: 870-878. https://doi.org/10.1016/j.ecolind.2015.08.036
Zeng P, Song X, Yang H, Wei N, Du L. 2022. Digital Soil Mapping of Soil Organic Matter with Deep Learning Algorithms. ISPRS Int J Geoinf, 11(5): 280-299. https://doi.org/10.3390/ijgi11050299
Zhang L, Heuvelink GB, Mulder VL, Chen S, Deng X, Yang L. 2024. Using process-oriented model output to enhance machine learning-based soil organic carbon prediction in space and time. Sci Total Environ, 922: 170778. https://doi.org/10.1016/j.scitotenv.2024.170778